



THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Motor activation in language processing:
Effects of handedness, experience, and planning

Madeleine E. L. Beveridge



PhD

University of Edinburgh

2014

DECLARATION

I hereby declare that this thesis is of my own composition, and that it contains no material previously submitted for the award of any other degree. The work reported in this thesis has been executed by myself, except where due acknowledgement is made in the text.

Madeleine E. L. Beveridge

Edinburgh, 1st May, 2014

ABSTRACT

Embodied Cognition accounts propose that motor activation contributes to semantic representations in action language (Fischer & Zwaan, 2008). However, the nature of this activation remains largely unspecified: in particular, which processes result in relevant activation? Long-term motor experience (e.g., the comprehender's dominant hand), short-term motor experience (e.g., the hand the comprehender has recently used), and action planning (e.g., the hand the comprehender is planning to use) are all potential candidates. This thesis uses a range of psycholinguistic methods (e.g., timed sentence-picture matching, two-alternative forced-choice sentence-picture matching, spoken sensibility judgements) to distinguish between these possibilities.

A first set of experiments investigated how comprehenders' handedness affects their interpretation of sentences describing manual actions (e.g., *I am slicing the tomato*). Participants matched sentences of actions to pictures of that action. The Body-Specificity Hypothesis (Casasanto, 2009; Willems, Hagoort, & Casasanto, 2010) predicts that right-handed and left-handed comprehenders will interpret manual action sentences differently, according to whether they would perform that action with their right or their left hand. However, we found that comprehenders appear to interpret manual action sentences according to the hand they use to respond to the task, and not the hand they would typically use to perform manual actions. In addition, this effect was stronger for first-person than third-person sentences, implying that the effect of motor activation is moderated by linguistic context.

A second set of experiments used the same paradigm but manipulated at what point comprehenders knew which hand they would use to respond to the sentences: during sentence processing, or after sentence processing was complete. We replicated the finding that comprehenders interpret manual action sentences according to their response hand, and that this effect was stronger for first- than for third-person sentences; but only when

comprehenders knew their response hand during sentence processing. In both sets of experiments, there was no effect of whether the picture of the action was presented from an egocentric or allocentric perspective, implying that action sentences are encoded for what effector (in this case, hand) will be used in the action, but not necessarily from what perspective the action will occur.

A third set of experiments investigated the existence of a causal role of action planning-based activation on sentence processing. Many studies have shown an effect of language processing on action execution (e.g., Glenberg & Kaschak, 2002; Glenberg et al., 2008), but a fully embodied theory of language also predicts an effect of motor activation on language processing. Here, right-handed participants made spoken judgements about sentences while planning an action with their right or left hand that matched or did not match the action described in the sentence. An effect of response hand on accuracy was found when the task required participants to explicitly judge the congruency of sentence and the action they were preparing, but not otherwise. These results corroborate recent research suggesting that activation of embodied lexical representations may be goal-driven rather than an automatic aspect of language processing (Hoedemaker & Gordon, 2013).

Overall, the experiments presented in this thesis suggest a possible role for planning-based motor activation in sentence processing, in line with embodied approaches; however, the results challenge strong accounts of embodiment by suggesting that the effect of planning-based activation is not automatic, and is moderated by linguistic context and task demands.

ACKNOWLEDGEMENTS

Many people have assisted me throughout the course of my PhD. First, I would to thank my supervisor, Martin Pickering, whose expertise and insights have been instrumental in shaping the thesis. His guidance as I attempted to steer a course through the vast and often conflicting literature in this field has been immeasurably useful. I'm extremely grateful to my second supervisor, Richard Shillcock, for his support and good humour throughout the doctoral process (there and back again). I would like to thank Antje Nuthmann, Martin Corley, Mateo Obregon and Ian Finlayson for their help and patience as I got to grips with programming and statistics. Daniel Casasanto and Roberto Bottini were collaborators on several studies, and provided valuable input into the experimental design and materials. I owe a huge thank you to Eddie Dubourg and Ziggy Campbell, whose technical expertise made possible the use of dual button boxes. More generally, the psycholinguistics community at Edinburgh has been a source of great support throughout my time here.

The University of Edinburgh generously funded my research through a Principal's Career Development Award, and also awarded me several research support grants. These grants allowed me to attend the 2011 Psycholinguistic in Flanders conference in Nijmegen, the 2011 Embodied and Situated Language processing conference in Newcastle, and the 2011 AMLAP conference in Riva del Garda. I am also very grateful to the Experimental Psychology Society, whose award of a Grindley grant allowed me to present a talk at the 2012 CUNY conference in South Carolina.

Finally, I'd like to say a huge thank you to my friends and family for their support over the past three years. Nat, Chez, mum and Carl – I am so very proud of us all. The usual online suspects (#magirlz) have made seemingly endless nights working late in the office pass much more quickly. And, last but not least, a special thank you to Oliver Davies, in return for countless cups of tea, cooked meals, wake up calls, and generally ensuring I stayed (a) alive, and (b) sane while writing up. You are indeed a very good egg.

TABLE OF CONTENTS

| | |
|---|----------|
| DECLARATION | III |
| ABSTRACT | V |
| ACKNOWLEDGEMENTS | VII |
| LIST OF TABLES..... | XV |
| LIST OF FIGURES..... | XIX |
| 1. INTRODUCTION AND OVERVIEW | 1 |
| 1.1. FOCUS OF THE THESIS | 1 |
| 1.2. CHAPTER BY CHAPTER OVERVIEW | 4 |
| 1.3. COLLABORATIONS AND PRESENTATIONS | 6 |
| 2. LITERATURE REVIEW | 9 |
| 2.1. INTRODUCTION | 9 |
| 2.2. MOTOR RESONANCE IN PERCEPTION | 10 |
| 2.2.1. <i>Mirror matching</i> | 12 |
| 2.2.2. <i>Effects of perception on action</i> | 19 |
| 2.2.3. <i>Effects of action on perception</i> | 24 |
| 2.2.4. <i>The role of experience</i> | 29 |
| 2.2.5. <i>Representations of self and other</i> | 32 |
| 2.3. MOTOR RESONANCE IN LANGUAGE | 37 |
| 2.3.1. <i>Language comprehension as action perception</i> | 37 |
| 2.3.2. <i>Types of embodiment theory</i> | 40 |
| 2.3.3. <i>Neural evidence for shared resources</i> | 43 |
| 2.3.4. <i>Effects of language on action</i> | 49 |
| 2.3.5. <i>Effects of action on language</i> | 56 |

| | |
|---|------------|
| 2.4. FLEXIBILITY IN EMBODIED REPRESENTATIONS | 62 |
| 2.4.1. <i>Sensitivity to context</i> | 63 |
| 2.4.2. <i>Flexibility in simulation: Prediction versus experience</i> | 70 |
| 2.4.3. <i>Body-Specificity</i> | 71 |
| 2.5. SUMMARY AND RESEARCH QUESTIONS | 76 |
| 3. PERSPECTIVE TAKING IN LANGUAGE: INTEGRATING THE SPATIAL AND ACTION DOMAINS..... | 77 |
| 3.1. INTRODUCTION | 77 |
| 3.2. REPRESENTING OTHER PEOPLE’S ACTIONS | 79 |
| 3.3. A TAXONOMY OF PERSPECTIVE | 80 |
| 3.4. DO LANGUAGE USERS CONSISTENTLY ADOPT THE AGENT’S PERSPECTIVE? | 84 |
| 3.5. SPATIAL PERSPECTIVE-TAKING | 91 |
| 3.6. SITUATION MODELS: LINKING SPATIAL- AND ACTION- PERSPECTIVES | 93 |
| 3.7. SUMMARY | 100 |
| 4. STATISTICAL ANALYSES USED IN THIS THESIS..... | 103 |
| 4.1. MIXED-EFFECTS MODELLING..... | 103 |
| 4.2. MAXIMAL RANDOM EFFECT STRUCTURE | 103 |
| 4.3. MODEL CONSTRUCTION..... | 104 |
| 4.4. POWER ANALYSIS | 105 |
| 5. BODY-SPECIFICITY IN ACTION SENTENCES | 107 |
| 5.1. OVERVIEW OF THE CHAPTER | 107 |
| 5.2. INTRODUCTION | 108 |
| 5.2.1. <i>Self and other in action language</i> | 109 |
| 5.3. THE SENTENCE-PICTURE MATCHING PARADIGM | 112 |
| 5.4. EXPERIMENT 1 | 114 |
| 5.4.1. <i>Method</i> | 115 |

| | |
|--|------------|
| 5.4.2. Results | 119 |
| 5.4.3. Discussion | 122 |
| 5.5. EXPERIMENT 2A | 124 |
| 5.5.1. Method..... | 124 |
| 5.5.2. Results | 127 |
| 5.5.3. Discussion | 135 |
| 5.6. EXPERIMENT 2B | 136 |
| 5.6.1. Method..... | 137 |
| 5.6.2. Results | 137 |
| 5.6.3. Discussion | 145 |
| 5.7. EXPERIMENT 3 | 146 |
| 5.7.1. Method..... | 147 |
| 5.7.2. Results | 148 |
| 5.7.3. Discussion | 150 |
| 5.8. GENERAL DISCUSSION | 151 |
| 5.8.1. Sentence pronoun | 152 |
| 5.8.2. Image orientation | 154 |
| 5.8.3. Current motor context | 155 |
| 5.8.4. Demand characteristics..... | 156 |
| 5.9. SUMMARY | 157 |
| 6. DEFINING THE CURRENT MOTOR CONTEXT | 159 |
| 6.1. OVERVIEW OF THE CHAPTER | 159 |
| 6.2. INTRODUCTION | 160 |
| 6.2.1. Temporal dynamics and motor resonance..... | 161 |
| 6.2.2. Motor planning versus short-term motor experience | 163 |
| 6.3. EXPERIMENT 4 | 166 |

| | |
|--|------------|
| 6.3.1. <i>Method</i> | 166 |
| 6.3.2. <i>Results</i> | 168 |
| 6.3.3. <i>Discussion</i> | 174 |
| 6.4. EXPERIMENT 5 | 175 |
| 6.4.1. <i>Method</i> | 176 |
| 6.4.2. <i>Results</i> | 176 |
| 6.4.3. <i>Discussion</i> | 183 |
| 6.5. EXPERIMENT 6 | 185 |
| 6.5.1. <i>Method</i> | 186 |
| 6.5.2. <i>Results</i> | 186 |
| 6.5.3. <i>Discussion</i> | 192 |
| 6.6. GENERAL DISCUSSION | 194 |
| 6.6.1. <i>Prospective versus retrospective simulation</i> | 195 |
| 6.6.2. <i>Is dominant hand defined by hand-in-use?</i> | 198 |
| 6.6.3. <i>Image-screen congruency</i> | 199 |
| 6.6.4. <i>Strategic processing</i> | 200 |
| 6.7. SUMMARY | 201 |
| 7. A FUNCTIONAL ROLE FOR ACTION PLANNING? | 203 |
| 7.1. OVERVIEW OF THE CHAPTER | 203 |
| 7.2. INTRODUCTION | 204 |
| 7.2.1. <i>Measuring language processing</i> | 204 |
| 7.2.2. <i>Evidence for an effect of action on language</i> | 205 |
| 7.2.3. <i>Long-term motor experience, revisited</i> | 207 |
| 7.2.4. <i>Sentence sensibility judgements</i> | 208 |
| 7.3. EXPERIMENT 7 | 210 |
| 7.3.1. <i>Method</i> | 211 |

| | |
|---|------------|
| 7.3.2. <i>Results</i> | 214 |
| 7.3.3. <i>Discussion</i> | 219 |
| 7.4. EXPERIMENT 8 | 220 |
| 7.4.1. <i>Method</i> | 221 |
| 7.4.2. <i>Results</i> | 221 |
| 7.4.3. <i>Discussion</i> | 225 |
| 7.5. COMBINED ANALYSES FROM EXPERIMENTS 7 AND 8 | 226 |
| 7.5.1. <i>Results</i> | 227 |
| 7.5.2. <i>Discussion</i> | 228 |
| 7.6. EXPERIMENT 9 | 229 |
| 7.6.1. <i>Method</i> | 230 |
| 7.6.2. <i>Results</i> | 231 |
| 7.6.3. <i>Discussion</i> | 235 |
| 7.7. GENERAL DISCUSSION | 237 |
| 7.7.1. <i>Effects of planned action</i> | 237 |
| 7.7.2. <i>Effects of response hand</i> | 239 |
| 7.8. SUMMARY | 240 |
| 8. CONCLUSIONS..... | 243 |
| 8.1 THESIS AIMS..... | 243 |
| 8.2. SUMMARY OF RESULTS | 243 |
| 8.3. IMPLICATIONS AND FUTURE RESEARCH..... | 245 |
| 8.3.1. <i>Perspective-taking in action language</i> | 246 |
| 8.3.2. <i>Spatial grounding and good enough representations</i> | 251 |
| 8.3.3. <i>Action planning and affordances</i> | 254 |
| 8.3.4. <i>Implications for naturalistic language processing</i> | 255 |
| 8.3.5. <i>Motor resonance and prediction</i> | 258 |

| | |
|--|------------|
| 8.3.6. <i>A functional role for motor resonance?</i> | 260 |
| 8.4. CONCLUDING REMARKS | 264 |
| REFERENCES | 267 |
| APPENDIX A: THE EDINBURGH HANDEDNESS INVENTORY | 297 |
| APPENDIX B: MATERIALS USED IN EXPERIMENTS 1 – 6 | 301 |
| APPENDIX C: MATERIALS USED IN EXPERIMENTS 7 – 9 | 309 |

LIST OF TABLES

| | |
|---|------------|
| <i>Table 2-1. Summary of effects from the embodied action semantics literature</i> | <i>67</i> |
| <i>Table 5-1. Mean winsorised button release and button press RTs (ms) by condition in Experiment 1</i> | <i>120</i> |
| <i>Table 5-2. Predictors of button release and button press RTs in Experiment 1.....</i> | <i>121</i> |
| <i>Table 5-3. Frequency of right-handed and left-handed image choice by image-screen congruency in Experiment 2a.</i> | <i>130</i> |
| <i>Table 5-4. Frequency of right-handed and left-handed image choice by handedness, pronoun, and image orientation in Experiment 2a.....</i> | <i>131</i> |
| <i>Table 5-5. Predictors of image choice in Experiment 2a.....</i> | <i>132</i> |
| <i>Table 5-6. Mean winsorised button release and button press RTs (ms) by image-screen congruency in Experiment 2a</i> | <i>133</i> |
| <i>Table 5-7. Mean winsorised button release and button press RTs (ms) by response hand, pronoun, and image orientation in Experiment 2a.....</i> | <i>134</i> |
| <i>Table 5-8. Frequency of right-handed and left-handed image choice by image-screen congruency in Experiment 2b</i> | <i>140</i> |
| <i>Table 5-9. Frequency of right-handed and left-handed image choice by pronoun and image orientation in Experiment 2b.....</i> | <i>140</i> |
| <i>Table 5-10. Predictors of image choice in Experiment 2b.....</i> | <i>140</i> |
| <i>Table 5-11. Mean winsorised button release and button press RTs (ms) by image-screen congruency in Experiment 2b</i> | <i>141</i> |
| <i>Table 5-12. Mean winsorised button release and button press RTs (ms) by condition in Experiment 2b</i> | <i>142</i> |
| <i>Table 5-13. Predictors of image choice in Experiments 2a and 2b</i> | <i>144</i> |
| <i>Table 5-14. Mean winsorised Voice Onset Times (ms) by condition in Experiment 3</i> | <i>149</i> |
| <i>Table 5-15. Predictors of Voice Onset Times in Experiment 3</i> | <i>150</i> |
| <i>Table 6-1. Frequency of right-handed and left-handed image choice by image-screen congruency in Experiment 4</i> | <i>171</i> |

| | |
|--|-----|
| Table 6-2. Frequency of right-handed and left-handed image choice by response hand, pronoun, and image orientation in Experiment 4..... | 171 |
| Table 6-3. Predictors of image choice in Experiment 4..... | 172 |
| Table 6-4. Mean winsorised button press RTs (ms) by image-screen congruency in Experiment 4.. | 173 |
| Table 6-5. Mean winsorised button press RTs (ms) by response hand, pronoun and image orientation in Experiment 4 | 173 |
| Table 6-6. Frequency of right-handed and left-handed image choices by image-screen congruency in Experiment 5 | 178 |
| Table 6-7. Frequency of right-handed and left-handed image choice by response hand, pronoun, and image orientation in Experiment 5..... | 179 |
| Table 6-8. Predictors of image choice in Experiment 5..... | 180 |
| Table 6-9. Predictors of image choice in Experiment 5 by block | 181 |
| Table 6-10. Mean winsorised button press RTs (ms) by image-screen congruency in Experiment 5 | 182 |
| Table 6-11. Mean winsorised button press RTs (ms) by response hand, pronoun, and image orientation in Experiment 5..... | 182 |
| Table 6-12. Predictors of image choice across Experiments 4 and 5..... | 184 |
| Table 6-13. Frequency of right-handed and left-handed image choices by image-screen congruency in Experiment 6 | 188 |
| Table 6-14. Frequency of right-handed and left-handed image responses by response hand, pronoun, and image orientation in Experiment 6..... | 189 |
| Table 6-15. Predictors of image choice in Experiment 6..... | 189 |
| Table 6-16. Predictors of image choice in Experiment 6 by response hand | 191 |
| Table 6-17. Mean winsorised button press RTs (ms) by image-screen congruency in Experiment 6 | 192 |
| Table 6-18. Mean winsorised button press RTs (ms) by response hand, pronoun, and image orientation in Experiment 6..... | 193 |
| Table 7-1. Accuracy rates for sensibility judgments in Experiment 7..... | 215 |
| Table 7-2. Mean winsorised Voice Onset Times (ms) by condition in Experiment 7 | 217 |
| Table 7-3. Predictors of Voice Onset Times by condition in Experiment 7 | 217 |

| | |
|--|-----|
| Table 7-4. Mean winsorised button press RTs (ms) by condition in Experiment 7 | 218 |
| Table 7-5. Predictors of button press RTs by condition in Experiment 7 | 218 |
| Table 7-6. Accuracy rates for sensibility judgments in Experiment 7..... | 222 |
| Table 7-7. Mean winsorised Voice Onset Times (ms) by condition in Experiment 8..... | 223 |
| Table 7-8. Predictors for Voice Onset Times in Experiment 8..... | 224 |
| Table 7-9. Mean winsorised button press RTs (ms) by condition in Experiment 8 | 225 |
| Table 7-10. Predictors for button press RTs in Experiment 8..... | 225 |
| Table 7-11. Accuracy rates for sensibility judgments in Experiment 9..... | 232 |
| Table 7-12. Mean winsorised Voice Onset Times (ms) by condition in Experiment 9 | 233 |
| Table 7-13. Predictors for Voice Onset Times and button press RTs in Experiment 9 | 234 |
| Table 7-14. Mean winsorised button press RTs (ms) by condition in Experiment 9. | 235 |

LIST OF FIGURES

| | |
|--|------------|
| <i>Figure 5-1. Example stimuli in critical and filler conditions in Experiment 1</i> | <i>117</i> |
| <i>Figure 5-2. An experimental item in the four conditions in Experiment 2a.....</i> | <i>128</i> |
| <i>Figure 7-1. Example stimuli in critical and filler conditions in Experiment 7</i> | <i>214</i> |

1. INTRODUCTION AND OVERVIEW

1.1. Focus of the thesis

How do we understand a sentence such as *I am slicing the tomato*? Psycholinguists have debated fiercely whether levels of processing *within* language (e.g., semantics versus phonology in lexical access) are independent stages in processing (Frazier, 1987; Levelt, Roelofs, & Meyer, 1999), or form part of an interactive system in which one level of processing can interact with other levels (Dell, 1986; Macdonald, Paerlmutter, & Seidenberg, 1994). However, language itself has typically been conceived as a self-contained cognitive module involving computations over amodal symbols (Fodor, 1983; Pylyshyn, 1984).

In recent years however, this conception of an independent and amodal language system has been increasingly challenged by embodied approaches to cognition in general, and language in particular (Barsalou, 1999; Glenberg & Gallese, 2012; Pulvermüller & Fadiga, 2010; Zwaan & Taylor, 2006). These *embodied cognition* accounts of language processing have provided vast quantities of empirical data indicating that the language system is less independent, and less amodal, than had been assumed. However, the theories have arguably been less successful in providing positive evidence that allows us to specify in detail the mechanisms and constraints under which embodied approaches to language might operate. In this thesis, I report 9 experiments which aim to answer the following questions about embodied language processing: (1) from whose perspective (e.g., agent, observer) do comprehenders interpret action language?; (2) what are the contributions of long-term motor experience, short-term motor experience, and action planning, to the way comprehenders interpret action language?; and (3) does activation of the motor system play a causal role in action language comprehension?

Philosophers and cognitive scientists have provided various definitions of embodied cognition (Clark, 1999; Gibbs, 2006; Varela, Thompson, & Rosch, 1991; Wilson, 2002), but

what all such accounts have in common is a rejection of amodal symbols as the basis for cognition (Barsalou, 2008). In traditional cognitive science, perceptual input is somehow translated into amodal representations (a “language of thought”; Fodor, 1975), which are stored and manipulated outside of the sensorimotor systems (Barsalou, Simmons, Barbey, & Wilson, 2003). In embodied approaches such as Perceptual Symbol Systems (Barsalou, 1999), these amodal symbols are absent, and cognitive processing instead involves *perceptual symbols*: modal representations of concepts in their original perceptual coding. These perceptual symbols are then combined and reactivated to produce simulations of that perceptual or motor experience. The little understood process of translating modal input to amodal symbol is therefore no longer necessary. Note that simulation, in this context, refers to the *effortless* and *unconscious* recruiting of sensorimotor systems (Jeannerod, 2006), rather than an effortful and deliberate state of imagining such as might be present in mental imagery (see e.g., Willems, Toni, Hagoort, & Casasanto, 2010 for differences between action verb understanding and mental imagery).

One consequence of rejecting amodal symbols, is that embodied accounts also reject the cognitive “sandwich” (Hurley, 2001) and, as a result, the modularity principle. Cognitive science has traditionally ignored the possibility of interaction between perception, action, and “higher-level” cognitive processes, because of the view that cognition constitutes the important “meat” between the input of perception, and the output of action (Shapiro, 2010). Crucially, in this traditional modular view, the perception and action systems do not contribute to cognitive processes such as language comprehension. However, increasing amounts of evidence suggest that language processing interacts with the perceptual and motor systems in a way that goes beyond this perception-cognition-action scheme (Willems & Francken, 2012). For example, comprehenders are typically faster to respond to images that match the orientation or shape that was implied in an earlier sentence, suggesting that language comprehension involves a perceptual simulation of the described events (for a review, see Zwaan & Madden, 2005). Such evidence, coupled with emerging evidence that

language is processed in distributed neural networks rather than being restricted to language-only structures (e.g., Pulvermüller, 2005) challenges the traditional view of language processing as modular and amodal.

In response to this growing body of empirical work, many researchers now reject the strict delineation between language and action (Gallese & Lakoff, 2005; Glenberg & Robertson, 2000; Pecher & Zwaan, 2005; Willems & Hagoort, 2007). These embodied approaches to language are often treated as a single cohort, placed in opposition to modular, amodal accounts of language. However, researchers are becoming increasingly aware that this simple dichotomy between *embodied* and *non-embodied* language is not particularly useful: rather, we should be exploring how, and under what circumstances, the language and action systems interact with one another (Willems & Francken, 2012). Therefore, the aim of this thesis is not to prove or disprove embodied accounts of language processing. Instead, I aim to establish some of the constraints that govern the language-action interaction, notably: the perspective adopted by the comprehender (Chapter 5), motor planning versus motor experience as a source of motor resonance (Chapter 6), and the potential for a causal relationship between action and language (Chapter 7).

The activation of the motor system in response to perceptual stimuli is known as motor resonance (Rizzolatti, Fogassi, & Gallese, 2001). However, since language is a referential system, there are in fact two levels at which this resonance could occur: communicative motor resonance – that is, activation in response to the articulatory gestures involved in the speech act; and referential motor resonance – that is, activation in response to the actions described in a particular utterance (Fischer & Zwaan, 2008). This distinction is also made in discussions of “vehicle” (form) and “content” (semantic meaning) levels of embodiment (Gallese, 2008; Pickering & Garrod, 2009). When Pickering and Garrod (2007, 2013) argue that language comprehenders recruit the production system to predict what the speaker will say next, they are speaking in terms of *communicative* motor resonance (see also the motor theory of speech perception, Liberman & Mattingly, 1985; Liberman, Cooper,

Shankweiler, & Studdert-Kennedy, 1967; D'Ausilio, Craighero, & Fadiga, 2012). However, when researchers investigate the potential for interaction between action and semantics, they are speaking in terms of *referential* motor resonance. In the current thesis therefore, we use the term motor resonance in the referential, rather than communicative, sense. How these two forms of motor resonance (communicative and referential) might combine in language comprehension is an avenue that should be explored in future research (see Chapter 8), but is beyond the scope of the thesis.

1.2. Chapter by chapter overview

The thesis is concerned with the question, how do comprehenders understand sentences describing simple actions, such as *I am slicing the tomato*? This question is broken down into specific research questions, with each research question being addressed in a different empirical chapter. In Chapter 2, I review the literature. I provide an overview of motor resonance in action observation, outlining issues concerning the distinction between representations of self and other. This distinction will be important when investigating what perspective (e.g., agent or observer) comprehenders adopt in action language. I then review the evidence for interaction between the language and action systems, covering neurophysiological evidence for shared neural resources, as well as behavioural evidence and patient studies suggesting that the two systems interact with one another at a causal level. I end the chapter by addressing the instability of results in this field, and discussing the flexibility of embodied language representations, and the possibility that motor resonance might help us predict future actions.

Chapter 3 is a theoretical development, in which I bring together research into spatial perspective taking, and research into embodied perspective taking in action language. I discuss the different action perspectives available to comprehenders, and the situations in which comprehenders appear to adopt one perspective over another. I suggest that

comprehenders can only run a full action simulation when they have sufficient spatial grounding, and I discuss the role of situation models in providing this grounding.

Chapter 4 provides an overview of the analysis methods used throughout the empirical chapters. I provide a brief summary of linear mixed effects models, as implemented in R, and motivate the use of maximal random effect structure wherever this results in a converged model (Barr, Levy, Scheepers, & Tily, 2013).

In Chapter 5, I present 4 experiments in which I empirically address the question of whose perspective (agent versus observer) comprehenders adopt, by having left- and right-handed participants respond to action sentences presented in the first- and third-person. Results suggest that participants adopted an embodied agent's perspective, but that this perspective is more likely to be adopted for first-person sentences than for third-person sentences. The results from this chapter also suggest that this perspective is grounded in the current motor context of the task, rather than in the participants' long-term motor experience.

In Chapter 6, I present 3 experiments investigating whether the key factor in current motor context is short-term motor experience, or current motor plan. Results showed that the embodied effects only emerged when participants were able to form a motor plan during sentence processing, and showed no effect of short-term motor experience. When a motor plan was present during sentence processing, the results also replicated the finding that comprehenders adopt an embodied agent's perspective, which is stronger for first-person sentences.

In Chapter 7, I report three experiments in which I tested whether there was any evidence that motor activation plays a causal role in online language comprehension. I found some evidence that a planned action affected accuracy in a sentence categorisation task, but only when the task explicitly probed the congruency between planned action and the sentence. No experiment showed any effect of planned action on speed of sentence processing.

In Chapter 8, I summarize the data presented in Chapters 5 – 7, and discuss the implications of these studies for embodied accounts of language processing, as well as directions for future research.

1.3. Collaborations and presentations

Chapter 3 is based on a theoretical article published with Martin Pickering, in *Frontiers in Human Neuroscience*. Apart from some minor modifications and an additional paragraph on pp. 101-102, the chapter is identical to the published paper. As a result of this, some of the introductory text in this chapter may briefly repeat ideas or summaries from Chapter 2. The published paper is part of a research topic on *Perspective Taking: Building a neurocognitive framework for integrating the “social” and the “spatial”*, and can be found at the following link:

- Beveridge, M.E.L. & Pickering, M.J. (2013) Perspective taking in language: Linking the spatial and action domains. *Frontiers in Human Neuroscience*, 7, 577.
doi:10.3389/fnhum.2013.00577.

The experiments in Chapters 5 and 6 are based on collaborative work with Daniel Casasanto and Roberto Bottini, and selected studies (notably Experiments 2a, 2b, 4, and 5) have been presented at the following conferences and workshops:

- Beveridge, M.E.L., Casasanto, D., Bottini, R., & Pickering, M.J. (2013). Motor plans and linguistic perspective-taking in sentence processing: A causal role in comprehension. Talk presented at the 26th CUNY Annual Conference on Human Sentence Processing, University of South Carolina.
- Beveridge, M.E.L., Casasanto, D., Bottini, R., & Pickering, M.J. (2012). Body-specificity in action sentences. Talk presented at the 4th Embodied and Situated Language Processing conference, Northumbria University, Newcastle-upon-Tyne.

- Beveridge, M.E.L., Casasanto, D., Bottini, R, & Pickering, M.J. (2012). Body-specificity in simple action sentences. Talk presented at the 11th Psycholinguistics in Flanders conference, Berg en Dal, Nijmegen.
- Beveridge, M.E.L., Casasanto, D., Bottini, R, & Pickering, M.J. (2012). Action sentences and body-specific representations. Poster presented at the 18th Architectures and Mechanisms of Language Processing conference, University of Trento, Riva del Garda.
- Beveridge, M.E.L., Casasanto, D., Bottini, R, & Pickering, M.J. (2012). How are action sentences represented? Poster presented at the 4th UK Cognitive Linguistics Conference, King's College London, London.

The experiments in Chapter 7 built on work carried out by Brian Gray for his master's thesis at the University of Edinburgh.

2. LITERATURE REVIEW

2.1. Introduction

Increasing amounts of evidence suggest some sort of link between language and action (Fischer & Zwaan, 2008). However, the nature of this link is not at all clear. Findings appear to vary depending on the precise task or the timing of stimuli, and theories rarely specify what type of effect (e.g., facilitation or interference) they predict. In order to move forward, researchers must move beyond demonstrating yet more embodied effects, but instead begin to investigate the constraints and contextual factors that appear to influence such effects (Willems & Francken, 2012). Only this way will we be able to begin to integrate the numerous empirical findings from embodied cognition, into a mechanistic and explanatory account of language comprehension. My aim in this thesis is to begin to specify some of these constraints. Specifically, I will investigate what perspective comprehenders adopt when interpreting action language sentences, and whether this perspective varies with the subject of the sentence (comparing first-person and third-person sentences). I will also investigate whether language processing interacts with motor planning, or motor experience. Finally, I will investigate a possible causal role of motor resonance on online sentence comprehension.

Before addressing these questions empirically, it is useful to gain an overview of the field as it stands. The recent upsurge in interest in embodied approaches to language grew out of findings from a literature apparently quite unconnected with language: action observation. The discovery of mirror neurons (see section 2.2.1) helped bring about a paradigm shift in which people began to view action and perception as related, rather than distinct systems (Rizzolatti & Craighero, 2004). As a result of this paradigm shift, researchers began to question whether other cognitive systems, such as language, might also

be more closely related to action than had been thought (Rizzolatti & Arbib, 1998). The literature on action perception therefore includes several parallels with the literature on action language comprehension; many of the debates and key questions concerning the former, also appear when investigating the latter. As a result of these parallels, I will begin the literature review by discussing some of the key findings and implications of the action perception literature (section 2.2).

Next, in section 2.3, I review the evidence for a link between language and action. I begin by discussing whether language can be seen as a form of action perception. I move on to discuss evidence from brain imaging that suggests distributed processing of language, and shared resources between the language and action systems. I end by outlining behavioural evidence for an action-language link. Here, I review both the evidence for an effect of language on action, including an overview of work on affordances; and the (much sparser) evidence for an effect of action on language.

Finally, in section 2.4, I discuss the instability that characterises much of the research into embodied approaches to language. I also note that for some researchers, the apparent flexibility of embodied representations suggests that simulations might help listeners to predict future actions. I therefore discuss whether simulation could be viewed as having a predictive component, as well as re-enacting previous experience.

2.2. Motor resonance in perception

Slicing a tomato and watching my friend slice a tomato appear to be clearly distinct events. In the first case, I know that it is me performing the action, and that I have control over how the action is performed. In the second case, I know that it is my friend performing the action, and that I am a mere bystander. However, over the past fifteen years, theories of action understanding have increasingly argued that the same mental representations are involved in both performing and in perceiving actions (e.g., Grèzes & Decety, 2001; Prinz & Hommel, 2002); the motor system appears to be activated during action observation as well as

execution. This recruitment of the motor system in tasks other than action planning and execution is known as motor resonance (e.g., Fischer & Zwaan, 2008). The finding of motor resonance has played a key role in shaping theories of motor cognition, which propose that action understanding involves simulating actions, either our own or other people's (Jeannerod, 2001; Jeannerod, 2006). These simulated actions are also known in the literature as covert actions (Jeannerod, 2006).

Research into motor resonance in perception has influenced embodied accounts of language processing in several ways. First, the fact that the motor system may be recruited for a task that is a step removed from performing an action (i.e., perceiving an action) raises the possibility that the motor system might also be recruited during other tasks removed from performing an action (i.e., understanding a linguistic description of an action). In particular, language may be understood through internal simulations (e.g., Barsalou, 1999) similar to the covert actions thought to occur during action perception (Jeannerod, 2001; Jeannerod, 2006). Second, although research has highlighted the overlap between self- and other-generated actions (motor resonance may allow us a “first person grasp” of other people's actions; Anquetil & Jeannerod, 2007; Rizzolatti & Sinigaglia, 2010), studies have also suggested that the motor system may be differentially involved in imagining self- versus other generated actions (Decety & Chaminade, 2003; Ruby & Decety, 2001). Therefore, if the motor system is involved in language processing, it may be differentially involved in processing self-referential language (e.g., *I am slicing the tomato*), compared with third-person language (e.g. *he is slicing the tomato*).

In the rest of this section, I therefore provide an overview of research into motor resonance during action perception (“*mirror-matching*”; Bouquet, Shipley, Capa, & Marshall, 2011; Schütz-Bosbach, Mancini, Aglioti, & Haggard, 2006). I begin by looking at the discovery that kick-started much of the research into the perception-action link: mirror neurons.

2.2.1. Mirror matching

Mirror neurons in non-human primates

In the 1990s, researchers discovered a group of neurons in the ventral premotor cortex (area F5) of macaque monkeys, which fire both when a monkey performs a particular action, and when the monkey observes that action being performed by another monkey or human (Di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). Unlike *canonical* neurons, which respond to the physical presence of an object, the *mirror* neurons discovered in area F5 respond specifically to interaction with an object (Jeannerod, Arbib, Rizzolatti, & Sakata, 1995; Rizzolatti & Luppino, 2001). The mere sight of an object does not lead to mirror neuron activity; neither does mimed action (e.g., making a grasping motion in the absence of an object) or simultaneously moving a hand and an object separately from one another (Di Pellegrino et al., 1992; Gallese et al., 1996; for a review of mirror neurons and their properties, see Rizzolatti & Craighero, 2004). Further research has suggested that mirror neurons might also exist in the inferior parietal lobe of the monkey brain (Fogassi et al., 2005).

Evidence for mirror-matching in humans

The discovery of mirror neurons in monkeys led researchers to hypothesise that a similar observation-execution matching system exists in humans. Mirror neurons were identified in monkeys (and more recently, song birds; Prather et al., 2008) using single cell recordings, however this technique has only recently been applied to human participants. Mukamel, Ekstrom, Kaplan, Iacoboni, and Fried (2010) recorded activity from single neurons in the medial frontal and temporal cortices of epilepsy patients as they observed or executed (a) hand movements or (b) facial expressions. A subset of these neurons (~7%) was found to fire when executing an action, and when observing that same action (i.e., the

neurons responded selectively to the observation and execution of either hand movements, or facial expressions). These findings constitute the first piece of direct evidence for the existence of mirror neurons in humans. However, they are far from conclusive. Although the findings suggest effector specificity in mirror-matching, they do not demonstrate the selectivity to particular movements or goals which is held to be a crucial part of the mirror neuron hypothesis (Dinstein, Thomas, Behrmann, & Haeger, 2008). Future work on single cell recordings in humans may provide further evidence for their existence, but at present, the ontological status of mirror neurons in humans is still debated (e.g., Hickok, 2009).

In the absence of other direct evidence for the existence of mirror neurons in humans, researchers have relied on brain imaging and electrophysiological techniques to infer the existence of a human mirror-matching system similar to that observed in monkeys. The lack of direct evidence for mirror neurons in human is, of course, problematic for researchers wishing to characterise the physiology of the human brain. However, indirect evidence such as that outlined below may in fact be more relevant than direct evidence in characterising the cognitive processes involved in action perception. For example, studies using electroencephalography (EGG) have demonstrated that action observation results in a similar inhibition of *mu* rhythm (typically present during rest, and inhibited during movement) as action execution (Altshuler, 2000; Cochin, Barthelemy, Lejeune, Roux, & Martineau, 1998; Cochin, Barthelemy, & Roux, 1999). Hari and Salmelin (1997) found that participants undergoing magnetoencephalography (MEG) experienced similar post-stimulus rebound effects for action observation as for action execution.

Fadiga, Fogassi, Pavesi and Rizzolatti (1995) stimulated the left motor cortex of participants using transcranial magnetic stimulation (TMS). Application of TMS results in the production, by affected muscles, of neuroelectrical signals called motor evoked potentials (MEPs). The amplitude of these MEPs is directly proportional to the level of activity in the motor cortex at that time (Papeo, Vallesi, Isaja, & Rumiati, 2009). Therefore, if observing an action involves motor activation, but observing an isolated object does not, then participants

should display different degrees of MEP activation in action observation versus object observation. Fadiga et al. tested this claim by applying TMS to the left motor cortex, and measuring the resulting MEPs from participants' hand muscles as they observed non-object-directed hand movements, object-directed hand movements, or stationary objects. Participants displayed increased MEPs relative to baseline for both types of movement, but not for object observation (see also Strafella & Paus, 2000).

Note that these results contrast with findings from non-human primates, in which non-object-directed movement typically does not trigger mirror neurons (Rizzolatti et al., 2001). Rizzolatti and Sinigaglia (2010) argue that whereas mirror neurons in monkeys respond to the goal of an action, the posited mirror system in humans can code either an action's goal, or the component movements leading to the accomplishment of that goal. The latter premise allows the possibility of activation triggered by movement without an object (and which therefore does not accomplish a goal), as observed by Fadiga et al. (1995). The proposal that mirror neurons in monkeys code for goals rather than lower level actions is supported by research demonstrating first, that the same neurons fire when performing actions using normal pliers, as when using modified pliers that require the opposite actions to achieve the same goal (Umiltà et al., 2008), and second, that different neurons fire in response to a "feeding" grasping movement than to a kinematically equivalent grasping movement with a different goal (Fogassi et al., 2005). However, the differences in methodologies employed by these researchers (i.e., indirect methods in humans versus direct methods in monkeys) make it difficult to draw firm conclusions about differences between the human and non-human systems.

Attempts to localise the mirror-matching system in humans using functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) have implicated the premotor and parietal cortices (e.g., Dinstein, Hasson, Rubin, & Heeger, 2007; Gazzola & Keysers, 2009; Grèzes, Armony, Rowe, & Passingham, 2003; Iacoboni et al., 2005). However, Dinstein et al. (2008) highlight two problems with attempts to localise the mirror

matching system. First, the use of fMRI and PET does not allow researchers to differentiate activation in *the same cells* (i.e., mirror neurons) during both observation and execution, from activation in *distinct neural populations in the same area*. In the original non-human primate studies, only a small population of cells in area F5 (~17%) were found to have mirror properties (Gallese et al., 1996). It therefore does not follow that any and all activation in homologue areas in humans during action observation and execution can be attributed to mirror neurons. Recent studies have begun using repetition suppression as a means of overcoming this criticism (Kilner, Neal, Weiskopf, Friston, & Frith, 2009).

However, a related problem is that most of the studies finding activation in the premotor and parietal cortices have, for theoretical reasons, looked for activation specifically in those areas. Thus, researchers presuppose the existence of a mirror-matching system in a particular area of the brain, and attribute any activation found in that area to this posited mirror-matching system (Dinstein et al., 2008). Similar arguments about the limitations of neural localisation emerge when trying to demonstrate that the language and action systems share neural resources (section 2.3.3), and highlight the importance of synthesising research from a range of methodologies, including behavioural experiments.

How specific is mirror-matching?

Interestingly, research suggests that activation of the motor system in action perception is not a generalised phenomenon. Instead, activation appears specific to the particular action being observed. This specificity is important, as it suggests that motor system involvement in language might code for the particular actions being described and perhaps also, the particular effector used to perform an action. In support of specificity, Buccino et al., (2001) found fMRI evidence suggesting that mirror-matching is sensitive to the effector involved (e.g., hand, mouth or foot). Participants observed video clips of object-directed movement, non-object-directed movement, or stationary objects. Both classes of movement, but not the stationary objects, triggered somatotopic activation in the premotor

cortex. In other words, clips of hand movements resulted in activation in areas associated with planning and executing hand movements, clips of foot movements resulted in activation in areas associated with planning and executing foot movements, and clips of mouth movements resulted in activation in areas associated with planning and executing mouth movements. Studies using TMS have suggested that mirror-matching is sensitive to not only the type of effector involved, but the laterality of that effector (e.g., left or right hand; Aziz-Zadeh, Maeda, Zaidel, Mazziotta, & Iacoboni, 2002), the effector orientation (e.g., same as or different to that of the participant; Maeda, Kleiner-Fisman, & Pascual-Leone, 2002), and the particular muscle involved in the task (e.g., handwriting versus arm movement; Strafella & Paus, 2000).

The finding that mirror-matching appears to code the particular effector, and even muscle, involved in the observed action, has led researchers to question whether mirror-matching extends to actions outside a participant's own motor repertoire. For example, monkeys show activation of mirror neurons when observing a human performing the same act (di Pellegrino et al., 1992), but what about when an animal observes actions that it is not capable of performing itself? Buccino et al., (2004) showed video clips of humans, monkeys, and dogs performing biting gestures, or communicative mouth gestures (speech mouthing, lip-smacking, barking, respectively) to human participants undergoing fMRI. While the biting movements all corresponded to actions within the human motor repertoire, the communicative gestures could be divided into those that are part of the human motor repertoire (i.e., speech mouthing, lip smacking), and those that are not (i.e., barking). Results showed that observing clips of biting movements elicited similar neuronal activity in the human participants, regardless of the species in the clip. However, of the communicative gestures, speech mouthing and lip-smacking produced similar activity to one another, while barking produced a distinct pattern of activity, in areas associated with visual, rather than motor, processing. This finding suggests that actions outside of an observer's motor

repertoire may be coded differently (e.g., in terms of perceptual features) than actions within an observer's motor repertoire (e.g., in terms of goals or motor processing).

The role of mirror neurons in action understanding

Early research on mirror neurons focused on the correspondence between executed and observed actions. However, more recent studies suggest that mirror neurons also fire in response to action-related sounds, such as peanut breaking or paper ripping (Keysers et al., 2003; Kohler et al., 2002; for a review of the link between actions and their sounds, see Aglioti & Pazzaglia, 2010). Umiltà et al. (2001) had monkeys observe actions in which the final part of the action (e.g., the point at which the experimenter's hand grasped an apple) was obscured. The authors found that a subset of previously identified mirror neurons discharged, even though the hand-object interaction was not visible to the monkey. The fact that mirror neurons can also fire in response to actions that are not fully observed suggests that the role of mirror neurons might not be restricted to aiding imitation, which would depend on the presence of visible hand-object interaction (e.g., Jeannerod, 1994). Mirror matching might instead be instrumental in action understanding (e.g., Rizzolatti et al., 2001; Rizzolatti & Sinigaglia, 2010; Umiltà et al., 2001). According to this view, mirror-matching allows an observer to understand an observed action by mapping it onto their motor repertoire (cf. Buccino et al., 2004). The observed action triggers the motor system to "resonate" (Rizzolatti, Fadiga, Fogassi, & Gallese, 1999), which allows the observer to infer the agent's goal and, therefore, understand the action (Blakemore & Frith, 2005; Fogassi et al., 2005; Rizzolatti et al., 2001). In this view, mirror-matching plays a causal role in action understanding, by allowing an observer to simulate performing the action herself, and thus understand the agent's goal (e.g., Jacob & Jeannerod, 2005; Jeannerod, 2006). Such a suggestion is of great interest to advocates of embodied approaches to language: if motor activation helps us understand other people's actions, even when they are not fully observed,

then a similar mechanism might also help us to understand *descriptions* about actions that are not observed.

However, Csibra (2007) proposes an alternative account, in which mirror-matching occurs after action understanding, *because* an observer already understands the intended goal of the agent (for a similar critique of the mirror neuron hypothesis as applied to social cognition, see e.g., Jacob, 2008, 2009). In this view, mirror-matching helps us predict and monitor an agent's future actions through a type of emulation (Prinz, 2006; see also Wilson & Knoblich, 2005), but it does not help us to understand the *current* action. Moreover, Hickok (2009) argues that the studies purporting to show that mirror-matching is involved in action understanding in fact show that activation of the mirror system is *not* necessary for action understanding. For example, Buccino et al., (2004) found activation of the motor system when participants observed human and chimp lip-smacking, but not dog barking. Presumably however, participants still understood the dog's action of barking – unless by “understanding”, Buccino and colleagues mean something above and beyond the usual meaning of the word. A precise definition of what this action understanding entails is therefore required, since a definition in which it is taken to mean activation of the motor system, is unsatisfactorily circular.

We will encounter these issues again in relation to action language understanding. For example, if motor simulation is taken as *constituting* action language comprehension, then the implication is that participants in experimental conditions where motor simulation is not present, do not understand the linguistic stimuli in those conditions. In section 2.3.2, I will discuss how Taylor and Zwaan's (2009, 2013) proposal of fault-tolerant language comprehension might help us resolve this issue. Similarly, the debate about the causal status of mirror-matching in action understanding, echoes a similarly fraught debate about the functional role of motor activation in action language understanding (sections 2.3.2. – 2.3.5). In Chapter 7, I will present 3 experiments that investigate whether there is any evidence of

such a functional role for motor resonance in online language comprehension. In the meantime however, I continue to review evidence for a link between action and perception.

2.2.2. Effects of perception on action

The research outlined above suggests that there may be a link between action and perception in humans, but it does not provide any evidence as to the direction of this relationship. In the following section, I outline evidence that observing an action leads to activation of the motor system: an effect of perception on action.

Explanations of the perception-action link

It is important to bear in mind that the research outlined below need not be interpreted as proof of a mirror neuron system in humans (see e.g., Hickok, 2009, 2010, for problems with this approach). Although the results are consistent with a human mirror neuron system, other theories make similar behavioural predictions, without committing to any particular claim about the neurological underpinnings of these predictions (Craigheero, Bello, Fadiga, & Rizzolatti, 2002). For example, motor cognition claims that people perform unconscious motor simulations (covert actions) of other people's actions, resulting in shared representations between self- and other-generated actions. Although recent formulations of this approach (e.g., Jeannerod, 2006) have drawn on mirror neurons as a possible mechanism through which motor simulations occur, they are not the only possibility. One could advocate the basic premise of motor cognition (simulations) without committing to the existence of specific mirror neurons in humans. Similarly, ideomotor theories of action such as Theory of Event Coding (TEC; Hommel, Müsseler, Aschersleben, & Prinz, 2001) stress the link between the internal image or perceptual consequences of an action, and the action itself, but remain agnostic as to the neural mechanism underpinning this link. Whether or not mirror neurons do underpin the link between action and perception is beyond the scope of this thesis. Similarly, although we will occasionally discuss the implications of TEC and motor

cognition for embodied approaches to language, we make no specific attempt to evaluate or distinguish between these different approaches to motor resonance (see, e.g., Jeannerod, 2006, p. 134-135 for differences between the two accounts).

Automatic imitation

Imitation in a social context has been widely studied: Chartrand and Bargh (1999) coined the term *Chameleon Effect* to capture the fact that in a social setting, people tend to unknowingly adopt the characteristics of the person they are engaging with. Such mimicry has been demonstrated, for example, for gestures and body movements (e.g., Bernieri, 1988; Shockley, Santana, & Fowler, 2003), and for facial expressions (Bavelas, Black, Lemery, & Mullet, 1986; Dimberg & Thunberg, 1998), even following unconscious perception of masked facial expressions (Dimberg, Thunberg, & Elmehed, 2000). In cognitive psychology, researchers have employed various terms to capture the finding that observing someone perform an action makes it more likely that the observer will perform that action herself: e.g., *automatic imitation* (Press, Bird, Flach, & Heyes, 2005), *visuomotor priming* (Craighero, Fadiga, Umiltà, & Rizzolatti, 1996), and *motor contagion* (Blakemore & Frith, 2005); see Heyes (2011) for a review. For clarity's sake, we use the term *motor resonance* to refer to activation of the motor system in non-motor tasks, generally, and *automatic imitation* to refer to the specific case of motor resonance during action observation.

Research has, over the past decade, provided ample evidence for automatic imitation. For example, people are faster to respond to clips of human body movement if the response (e.g., opening or closing their hand) matches the movement shown in the video (e.g., a hand opening or a hand closing), even when the observed movement is irrelevant to the task (Stürmer, Aschersleben, & Prinz, 2000). Similar results have been found in a go/ no go paradigm (Brass, Bekkering, & Prinz, 2001), using static images rather than video clips (Vogt, Taylor, & Hopkins, 2003), and using auditory rather than visual stimuli (Drost, Rieger, & Prinz, 2007; Haueisen & Knösche, 2001).

Whereas the studies outlined above showed a congruency advantage (actions are more likely to be performed, or performed more easily, having observed congruent actions), Kilner, Paulignan and Blakemore (2003) found the opposite effect. They studied the effect of perception on action by measuring variance in the kinematics of performed movements: participants executed arm movements while observing congruent or incongruent arm movements performed by a human or a robotic arm. Observing incongruent arm movements performed by a human arm led to significantly more variance in participant's own arm movements relative to baseline (see also Kilner, Hamilton, & Blakemore, 2007). Congruent arm movements and incongruent robotic arm movements had no effect. The difference in direction of effect between this and other studies (congruency *versus* incongruency advantage) might be attributed to task demands, in particular timing. In Kilner et al.'s study, participants performed and observed the actions at the same time; in the case of Brass et al. and Vogt et al., participants first observed, and then performed an action. However, in Drost et al.'s study, musicians heard and performed actions simultaneously, and yet participants showed a congruency advantage rather than an incongruency advantage. Inconsistency in effect direction is not restricted to the perception-action literature; it is also a notable feature of research into the interaction between action and language. Arriving at a satisfactory explanation of these inconsistencies, which will allow future research to predict the direction of effects, should be a high priority for theories of motor resonance in action perception and in language.

Automatic imitation in speech

Interestingly, an effect of perception on action has also been observed in speech. Fadiga, Craighero, Buccino and Rizzolatti (2002) used TMS to measure MEPs from Italian participants' tongue muscles as they listened to speech. Recall from section 2.2.1 that the greater the activation present in the motor cortex, the greater the amplitude of the MEPs in the corresponding muscle. Tongue MEPs increased significantly when participants listened

to words that featured a sound involving a strong tongue movement (e.g., *birra* [beer]) compared with words that did not (e.g., *baffo* [moustache]). In other words, participants showed evidence of increased activation in tongue motor areas when *perceiving* words that involved strong tongue movement, even though they did not *execute* any tongue actions themselves. Pulvermüller et al. (2006) found evidence for somatotopic activation of articulators (lips, tongue) as participants listened to speech while undergoing fMRI. And in a behavioural paradigm, Kerzel and Bekkering (2000) showed an effect of stimulus-response compatibility in speech. Participants observed a speaker pronouncing the syllables *ba* or *da*, while the printed word *ba* or *da* was superimposed on the speaker's lips to indicate what the participant should produce. Participants were faster to produce syllables that were congruent with the observed speech, even though the observed speech was irrelevant to the task. Together, these results lend some support to a motor theory of speech perception, in which comprehenders understand speech by covertly producing the articulatory gestures necessary to produce the sounds they are hearing (Galantucci, Fowler, & Turvey, 2006; Liberman & Mattingly, 1985; Liberman et al., 1967). However, the necessity of such activation for speech perception remains unproven (see Willems & Hagoort, 2007 for a review of relevant neural evidence).

The role of animacy

The findings of Kilner et al. (2003), in which observation of human, but not robotic arm movements affected participants' performance to perform movements themselves, suggest that the link between action execution and observation may be limited to observation of actions by biological entities. This possibility is supported by results from fMRI studies showing differential activation when observing actions performed by human hands versus 3D virtual reality hands (Perani et al., 2001), or robotic hands (Tai, Scherfler, Brooks, Sawamoto, & Castiello, 2004). However, the observed bias may involve more than just animacy: the bias may also involve how likely the observer is to empathise with the observed

entity. Crescentini, Mengotti, Grecucci and Rumiati (2011) adapted a paradigm used in a behavioural study by Brass et al. (2001), for use in an fMRI study. Participants performed a finger movement (e.g., tapping) having watched a biological (e.g., hand) or non-biological (e.g., dot) perform a compatible (e.g., tapping/ downwards movement) or incompatible movement (e.g., lifting/ upwards motion). Activation was recorded during the action execution, and showed a main effect of biological versus non-biological stimuli, a main effect of compatibility (different regions activated for incompatible versus compatible actions), and an interaction between these two following an emotionally neutral or sad (but not angry) context. The authors concluded that the compatibility effect for biological stimuli only occurred in situations where people are more likely to empathise, or take another's perspective.

Behavioural studies have provided mixed evidence for an *animacy bias* (Bird, Leighton, Press, & Heyes, 2007). For example, Castiello, Lusher, Morena, Edwards and Humphreys (2002) found that the effect of automatic imitation was stronger following observation of a human hand compared with observation of a robotic hand (see also Jonas et al., 2007). However, Press, Bird, Flach and Heyes (2005) had participants perform a pre-determined action (e.g., open their hand) in response to a congruent (e.g., open hand) or incongruent (e.g., closed hand) action by a human or a robotic hand. Although the effect of automatic imitation was stronger in the biological (human hand) condition than the non-biological (robotic hand) condition, participants were still faster in the congruent versus incongruent conditions for the robotic hand. Thus, the robotic hand still appeared to result in some automatic imitation. Similarly, Liepelt and Brass (2010) found that, although automatic imitation was stronger for observation of a human hand *versus* a wooden hand, automatic imitation still occurred in the wooden hand condition. However, in Liepelt and Brass's study, the actual observed movements were the same in both conditions. The only difference between conditions was that prior to testing, participants were shown a human hand or a wooden hand inserted into a glove, and told that the subsequent movements were

made by that hand. Therefore, in this study, it appears that the participants' belief about the biological status of the hand, rather than any visual input from the hand itself, affected the degree of automatic imitation.

The key to the biological/ non-biological difference may therefore lie not in the movement profile itself, but rather in a participant's belief about the potential agency of the actor. Stanley, Gowen and Miall (2007) showed participants clips of biological (produced by the experimenter) or non-biological dot motion. Following presentation of the movement clip, participants moved their arm in the same plane (congruent) or a different plane (incongruent) to that of the dot motion clip. Participants who were told that all the movements they observed were produced by a human showed an interference effect in their own movement execution. Participants who were told that the movements they observed were computer generated showed no interference effect. These results suggest that belief about the other's potential to act as an agent affects how people represent their action. When Press, Gillmeister and Heyes (2006) explicitly questioned participants in an automatic imitation task about the animacy of the stimuli, the reported belief of animacy had no effect on participants' performance. However, an implicit assessment of participant's animacy belief may provide a better test for the hypothesis that belief in animacy does not affect automatic priming (e.g., Press et al. note that participants may have been reluctant to use extreme scores on the animacy rating scale).

2.2.3. Effects of action on perception

The link between action and perception does not appear to be one-way. Schütz-Bosbach and Prinz (2007) review evidence for *perceptual resonance*: an effect of action on perception. Several studies appear to demonstrate that perceiving a stimulus while simultaneously planning and executing an action leads to a very generalised reduced perception of the visual stimulus. For example, De Jong (1993) found that participants are slower to correctly identify letters while concurrently listening, and planning to respond to

an auditory stimulus. However, it is unclear whether this effect is actually due to an effect of action on perception, rather than a more general effect where resources of a single processing channel are initially allocated to the first task, and thus unable to attend to the second, overlapping task (De Jong & Sweet, 1994; De Jong, 1995; Joliceœur, 1999; see Pashler, 1994, for a review).

Blindness to response-compatible stimuli

In order to demonstrate a specific effect of action on perception, researchers must provide evidence of more specific cross-talk between these two domains: the difference in perception must be shown to vary according to the relationship between the action and the perceived stimuli (Müsseler, Wühr, & Prinz, 2000). Blindness to response-compatible stimuli is one such effect. Müsseler and Hommel (1997a) had participants perform a left or right key press, just as they were presented with a left or right pointing masked arrow. Participants were less able to correctly identify the direction of the arrow (left or right) when it was congruent with the direction of their button press (left or right). These findings suggest that people find it harder to perceive a stimulus when they are concurrently carrying out an action that shares certain features (e.g., direction) with that stimulus than otherwise.

This particular effect is predicted by the Common Coding Approach (Prinz, 1990), in which stimuli and responses are represented in the same system. This system is one of external events, whereby stimuli represent perceived events, and responses represent planned events (see Theory of Event Coding, Hommel, Müsseler, Aschersleben & Prinz, 2001, for a later development of this approach). The Common Coding Approach posits that since both stimuli and response are represented within the same system, they can be coded for the same features. So, both a stimulus (such as an arrow), and a response (such as a button press), can share the code LEFT. This means that perceiving a left-pointing arrow would trigger a left-sided response. However, it also means that performing that left-sided response would, in turn, reactivate the perceptual code for a left-sided response. The question then arises: how

does the actor break this potentially endless cycle of perception triggering action triggering perception? Mackay (1987) argues for a self-inhibition phase, whereby once a feature code has been activated, it results in lower than normal activation of that code. In this way, as the actor executes a left-sided response, the perceptual system is rendered less sensitive to the code LEFT, and is therefore not triggered by the LEFT-ness of the response. The perception-action loop can thus be avoided. If such a self-inhibition phase exists, then blindness to response-compatible stimuli should occur when participants are asked to perceive a stimulus while already initiating a response that shares a feature code (e.g., LEFT) with that stimulus. In the case of Müsseler and Hommel (1997a), participants initiated a left button press, thus lowering the activation of the LEFT code, so that the left-pointing arrow was perceived less easily than an arrow that did not share the LEFT code (e.g., a right-pointing arrow).

Further work has replicated this blindness effect in a task where participants were required to detect, rather than identify the direction of, masked arrows during a left or right key press (Müsseler & Hommel, 1997b), and during tasks where the cue telling participants which action to perform (in the original studies this was also a left or right pointing arrow) shared only partial or no feature overlap with the perceived stimulus or the executed action (e.g., black or white squares; Müsseler, Steininger, & Wühr, 2001; Müsseler et al., 2000). Kunde and Wühr (2004) found that the blindness effect persisted independently of the hand used to execute the action, and that a similar effect is observed when producing colour words (action) and perceiving patches of colour. These results, they suggest, imply that blindness to response-compatible stimuli is a general phenomenon whenever the action and perceptual stimuli overlap conceptually.

Underlying the importance of the meaning attached to the perceptual stimulus, Stevanovski, Oriet and Jolicoeur (2002) found that the effects could be reversed if the participants interpreted the same visual stimulus as pointing in the opposite direction; for example, the direction of effects switched when participants interpreted the > symbol as a headlight (beaming light to the left) rather than an arrow (pointing to the right). However,

Hommel and Müsseler (2006) argued that semantic congruence is insufficient to elicit blindness to response compatible stimuli. They tested the generality of the blindness effect using combinations of left- or right-pointing arrows, and the words *left* or *right* as stimuli; and left- or right-button presses, and vocal responses *left* or *right* as responses. The arrows paired with the button presses produced the expected effect, as did the words paired with the vocal responses. But when the arrows were paired with the vocal responses, or the words with the button presses, the blindness effect did not occur. In other words, even though all the pairs overlapped conceptually, the effect only occurred in those pairs that also overlapped in the perceptually-derived features, in line with TEC accounts.

Facilitation effects

The effects outlined above constitute evidence of an incongruency advantage, or contrast effect of action on perception (Schütz-Bosbach & Prinz, 2007). Similar effects have also been found for more complex actions, such as lifting boxes and estimating the weight of boxes lifted by others (Hamilton, Wolpert, & Frith, 2004); performing arm movements and determining the direction of moving stimuli (Zwicker et al., 2008; Zwicker, Grosjean, & Prinz, 2010); and drawing sinusoidal lines without visual feedback, while observing a dot moving in sinusoidal trajectories of various amplitudes (Schubö, Aschersleben, & Prinz, 2001; Schubö, Prinz, & Aschersleben, 2004). However, as in the case of automatic imitation, other studies have found effects occurring in the opposite direction.

Facilitation, or assimilation effects occur when participants become more, rather than less sensitive to stimuli that match their action. For example, Wohlschläger (2000) had participants turn a knob clockwise or anticlockwise, and at the same time report the perceived motion direction of a circle of dots (pushing a right or left key to indicate clockwise or anticlockwise movement, respectively). The dots were in fact ambiguous as to which direction they were moving in, but participants were more likely to perceive the dots as moving in a direction that was congruent with their hand movement (clockwise or

anticlockwise). Similarly, Repp and Knoblich (2007) found that the direction of movements required to create a pair of ambiguous tones on a piano keyboard (right key followed by left key; left key followed by right key) influenced perceived changes in pitch in an ambiguous tone sequence; participants were more likely to hear pitch as rising using a left to right key press order than a right to left key press order (but only for expert piano players).

Other research suggests that the specific type of hand action being planned (e.g., grasping or pointing) can facilitate perception of congruent stimuli. Wykowska, Schubö and Hommel (2009) asked participants to prepare a grasping or a pointing action, and then perform a visual search task. They found that the type of action participants planned to make influenced detection of action-relevant features in the visual search task: participants were faster to detect different size targets when they prepared a grasping action, and faster to detect different luminescence targets when they prepared a pointing action (see also Bekkering & Neggers, 2002). Further studies have found effects of hand action on change-detection – participants planning a grasping movement were more sensitive to change detection when a grasp-congruent object changed between two otherwise identical pictures (Symes, Tucker, Ellis, Vainio, & Ottoboni, 2008) – and detecting oddballs in a sequence of images (Fagioli, Hommel, & Schubotz, 2007; Fagioli, Ferlazzo, & Hommel, 2007).

Note that the Common Coding Approach specifically predicts an incongruency advantage, since the activation of a particular feature code should leave the agent less sensitive to that code (Müsseler & Hommel, 1997a). To try and explain the differences in effect direction between studies, Zwickel, Grosjean and Prinz (2010b) note that studies finding contrast effects typically use unambiguous stimuli such as left- or right-pointing arrows, whereas studies finding assimilation effects typically use ambiguous stimuli such as ambiguous piano tones or ambiguous movement (e.g., Repp & Knoblich, 2007; Wohlschläger, 2000). It is possible, therefore, that the nature of the experimental stimuli (ambiguous versus unambiguous) affects whether the results show a congruency or incongruency advantage. Although change detection, oddball and visual search tasks do not

constitute ambiguous stimuli as such, these tasks do involve searching for a stimuli rather than observing it in an unambiguous context. The extent to which task demands might explain differences in effect direction is explored more fully in section 2.4.1.

A more general difficulty that researchers may face when investigating the perception-action link, is distinguishing the effects of action on perception from those of perception on action. For example, Craighero, Fadiga, Rizzolatti and Umiltà (1999) report a series of experiments showing that participants who prepared to grasp a bar in a particular orientation were faster to execute the grasping movement following presentation of congruent versus incongruent stimuli in a go/ no-go task (see also Craighero et al., 2002). Craighero et al. (1999) interpret these findings as showing that preparing to act leads to faster processing of visual stimuli – in other words, an effect of action on perception. However since the measured outcome variable is action, it is impossible to distinguish this explanation from one in which processing of congruent visual stimuli affects action execution – in other words, an effect of perception on action. Similarly, when discussing the link between language and action, the frequent use of motor actions as a response means that it is often unclear whether a given effect is due to activation in the motor system affecting language comprehension, or language comprehension affecting motor execution. This uncertainty is one reason why the functional role of the motor system in language comprehension is contested, and requires further investigation (see Chapter 7).

2.2.4. The role of experience

The Associative Hypothesis

Heyes (2010) notes that many discussions of mirror neurons implicitly assume that such neurons evolved through adaption to the demands of natural selection, and that as a result, many of their properties are genetically inherited (e.g., Rizzolatti & Arbib, 1998; Rizzolatti & Craighero, 2004). Although such an account cannot be ruled out, Heyes favours

an Associative Hypothesis, where mirror neurons are formed through sensorimotor experience. For example, a sensory neuron might fire in response to perceiving a grasping action, and another motor neuron might fire in response to performing a grasping action. Perceiving and performing actions is often correlated (e.g., when a parent imitates a child's action), meaning that the motor neuron will often fire when the grasping action is perceived, because it is also being performed. Eventually, through this association, the motor neuron will fire when observing that action, even when it is not performed (see Heyes, 2010 for details; Pulvermüller, 1999, offers a similar account of how words come to be associated with their meanings). The Associative Hypothesis can, Heyes argues, provide a better explanation of why mirror-matching in adults can be affected by learned sensorimotor associations based on short-term motor experience, compared with evolutionary accounts.

Short-term sensorimotor couplings affects motor resonance

In this thesis, I will explore the contributions of short-term and long-term motor experience on language processing. The idea that short-term motor experience might override the effect of long-term motor experience in language processing, is suggested by evidence that short-term sensorimotor experience can override the effect of long-term sensorimotor experience in automatic imitation. For example, we saw in section 2.2.1 that automatic imitation is typically elicited by human, but not robotic hands. However, 24 hours after a brief training period (216 trials) in which participants opened and closed their hands in response to open or closed robotic hand stimuli, respectively, the robotic hands elicited as much automatic imitation as the human hands (Press, Gillmeister, & Heyes, 2007). In another study, automatic imitation effects were removed by training participants to respond to open-hand stimuli by closing their hands, and to respond to close-hand stimuli by opening their hands (Heyes, Bird, Johnson, & Haggard, 2005; see Gillmeister, Catmur, Liepelt, Brass, & Heyes, 2008 for similar effects with hand and foot actions).

The suggestion that short-term sensorimotor experience modulates automatic imitation is supported by evidence from studies using TMS. Catmur, Walsh and Heyes (2007) measured MEPs in participants' index and little finger abductors, before and after training sessions in which participants responded to little finger movements by moving their index finger, and vice versa. Prior to training, participants displayed the expected mirror-matching effect: observing a little finger movement led to increased MEPs in little finger, but not index finger abductors; and observation of an index finger movement produced increased MEPs in index finger, but not little finger abductors. After training however, this effect was reversed. Observation of index finger movements led to increased MEPs in the little finger abductor, and observation of little finger movements led to increased MEPs in the index finger abductor. Together, these results suggest that short-term associations between perceptual input and performed actions can enhance, reduce, and even reverse more long-term associations that trigger motor resonance. One of the aims of this thesis therefore, is to investigate whether short-term experience can override the effect of long-term experience on action language comprehension.

The importance of motor cues

Sensorimotor experience involves associating a particular action with a particular sensory input (Catmur, Walsh, & Heyes, 2009). However, motor resonance might also be affected by sensory experience alone (e.g., how I observe most other people performing an action), or by motor experience alone (e.g., how I personally perform an action). Petroni, Baguear and Della-Maggiore (2010) found increased corticospinal activity when participants viewed abstract stimuli (coloured shapes) that had been associated with finger movements during a short (~12 minute) training session; importantly, this effect only occurred when stimuli had been motorically associated (participants performing action in response to the stimulus) rather than visually associated (participants viewing action after stimulus). These findings suggest that sensory experience may have a limited role in modulating motor

resonance, and are supported by fMRI research showing that expert dancers show greater bilateral activation when they view movements they have been trained to perform, compared with moves they view often but do not perform themselves (e.g., male and female ballet movements; Calvo-Merino, Grèzes, Glaser, Passingham, & Haggard, 2006). Similarly, Casile and Giese (2006) found that acquiring novel motor behaviours facilitated participants' perceptual discrimination of that action, even when they learnt the motor behaviour while blindfolded, thus eliminating visual feedback (see also Hecht, Vogt, & Prinz, 2001).

These results provide a possible framework for explaining the finding, discussed in section 2.2.5 below, that people tend to be better at recognising their own movements than those of other people, even though they have greater visual experience of other people's actions than of their own: motor cues have a stronger effect than visual cues in modulating motor resonance. The distinction between visual cues and motor cues will be important in discussing the possible perspectives that comprehenders might adopt on action language, as follows. If visual cues have a stronger role on motor resonance, then we might expect people to adopt the perspective of someone observing the described event. If motor cues have a stronger role on motor resonance, then we might expect people to adopt the perspective of someone performing the described event (see Chapter 3 for more details on the possible perspectives available to comprehenders).

2.2.5. Representations of self and other

Shared representations need not be identical

One consequence of the link between perception and action is the implication that we represent other people's (observed) actions in the same way as we represent our own (executed) actions. This is important because it suggests that comprehenders could interpret an action from someone else's point of view, even when they are describing or hearing about an action performed by somebody else. However, the hypothesis that our own and other

people's actions share representations leads to the problem of action attribution: how do we know whether a particular action was performed by ourselves or by somebody else? (de Vignemont & Haggard, 2008). The solution to this problem is that although self- and other-generated actions may share some representations, the representations cannot overlap completely, else self-generated and other-generated actions would indeed be indistinguishable. These non-overlapping aspects of the representations would then act as a cue to how we achieve agency attribution (Decety & Sommerville, 2003; Jeannerod, 2006).

Several studies suggest that self-generated actions may have a tighter link to the sensorimotor system than other-generated actions. For example, Grezes, Frith, Passingham (2004) had participants watch videos of themselves and others lifting a box. On 10% trials, the person in the video was told the incorrect weight of the box; participants judged whether or not they had been told the correct weight, while undergoing fMRI. Activity in the parietal premotor cortex began earlier when participants were making judgements about their own actions, than when they were judging other people's actions. In another fMRI study, Jackson, Meltzoff and Decety (2006) showed participants clips of simple hand or foot actions. The clips were either presented from the perspective of the agent (INTERNAL orientation; that is, the effector extended upwards from the bottom of the screen, as though it could plausibly be interpreted as belonging to the participant); or from the perspective of an observer (EXTERNAL orientation; that is, the effector descended downwards from the top of the screen, and could most plausibly be interpreted as belonging to an agent facing the participant). The internally orientated clips thus represented a first-person visual perspective, and the externally orientated clips represented a third-person visual perspective. Participants either watched the clips passively, or imitated the actions that they saw. Internally orientated actions elicited greater activation in the left sensorimotor area than externally orientated actions, even in passive observation.

The difference between perception of first-person and third-person actions raises the question of whether there might also be differences in the way that people understand

language about first-person or third-person actions. In the case of typical dialogue, we are used to interpreting first- and second-person language (*I am slicing the tomato, you are slicing the tomato*) deictically, such that the person assumed to perform a described action (self or other) depends on who is uttering the description (myself or my interlocutor). It may therefore be the case that we are trained to ignore person in the comprehension of action sentences, and that there would be no difference in the degree of motor activation in response to first- versus third-person sentences. On the other hand, third-person sentences represent an unambiguous case in which the described action is being clearly attributed to a third party; whereas first-person language can refer to either self (if I am speaking) or other (if my conversation partner is speaking). In both cases, *I* typically refers to someone taking part in the conversation. In a laboratory setting, when no conversation partner is present, comprehenders might interpret first-person language as referring to themselves, as the only available source of linguistic utterances.

If that is the case, then first-person action descriptions might result in increased motor activation relative to third-person action descriptions, mirroring the tighter connection between self-generated actions and the sensorimotor system. In support of this prediction, Papeo, Corradi-Dell'Acqua, and Rumiati (2009), using TMS, found increased MEPs in the hand muscles of participants reading first-person manual action verbs (e.g., *I grasp*) relative to first-person non-action verbs (e.g., *I think*), but no such difference between third-person action and non-action verbs. The authors argued that participants interpreted the agent of the first-person verbs to be themselves (in contrast to the third-person verbs, which they interpreted as referring to a third party), and that comprehenders simulate action language only when the self is interpreted as the agent of the action. I will test whether or not there appear to be behavioural differences in the interpretation of first- versus third-person action descriptions in Chapters 5 and 6. For more details on overlapping and non-overlapping cortical areas in self- and other- representations, see for example, Decety and Sommerville

(2003), Decety et al., (1994), Grafton, Arbib, Fadiga and Rizzolatti (1996), Ruby and Decety (2001), Ruby and Decety (2004).

Recognising self-generated actions

Despite the link between executing an action, and observing another person executing an action, people appear to have remarkably little difficulty recognising their own actions. For example, Knoblich and Prinz (2001) found that participants were able to distinguish their own drawings from those of other people, regardless of the familiarity of the symbol in the drawing, and despite receiving no visual feedback while drawing. People are also able to identify tapping noises made by themselves rather than by a computer (Knoblich & Repp, 2009), and recordings of themselves playing the piano, even if they received no auditory feedback at the time of playing (Repp & Knoblich, 2004); and to identify their own clapping, even when the actual claps are replaced by uniform tones (i.e., temporal information alone sufficed; Flach, Knoblich, & Prinz, 2004).

Not only can people recognise their own actions, but there is converging evidence that people are better at perceiving and predicting their own movements than those of other people. For example, Knoblich and Flach (2001) found that participants were more accurate at predicting the final location of a dart thrown at a dartboard when watching videos of themselves, rather than another person, throwing the dart. Note that even though such results stress the difference between self and other, they also support the link between action and perception, because such a link would predict that when the same system perceives and executes an action, it will make better predictions than when two different systems are involved (Knoblich & Flach, 2001). Similarly, Daprati, Wriessnegger and Lacquaniti (2007a) found that people were faster and more accurate at making same-different judgements in movements when pairs of movements were their own rather than those of another person, and Daprati, Wriessnegger and Lacquaniti (2007b) found that people were able to select their own action when a virtual hand performed motion-capture kinematics of

their own and other people's actions. The ability to recognise their own movements persisted even when participants received no visual input about their movements during the recording phase, implying that the results could be attributed to kinematics rather than visual feedback.

Loula, Prasad, Harber and Shiffrar (2005) found that participants were better at recognising their own dancing than that of other people in point light displays, despite the fact that we typically observe other people dancing more often than we observe ourselves. An advantage for self-generated actions has also been shown in gait (Beardsworth & Buckner, 1981; but see Cutting & Kozlowski, 1977), moving in time with music (Sevdalis & Keller, 2010), and lip reading (Tye-Murray, Spehar, Myerson, Hale, & Sommers, 2013). These results are important for two reasons. First, they support the hypothesis that self-generated actions are more tightly connected to the sensorimotor system than other-generated actions; and that, by extension, first-person descriptions of actions may be more connected to the sensorimotor system (more "embodied") than third-person descriptions. Second, they suggest that motor experience (how I perform a particular action) may play a greater role in action perception than visual experience (how other people perform a particular action): in the studies outlined above, participants were better able to recognise actions they perform themselves (increased motor cues), compared with actions performed by, and therefore observed in, other people (increased visual cues). If motor cues, rather than visual cues, have a privileged role in action perception, then they may also have a privileged role in interpreting action descriptions. In other words, comprehenders might be more likely to interpret a sentence such as *I am slicing the tomato* in line with how they personally would slice a tomato, rather than how they observe other people slice a tomato. In Chapter 5, I will investigate the role of motor versus visual experience in comprehending simple action sentences.

2.3. *Motor resonance in language*

In the previous section, I discussed the evidence for motor resonance during action perception. The evidence for some form of activation of the motor system during action perception is compelling, but the nature of the relationship between action execution and perception remains unclear. If we consider language comprehension to be a type of perception (e.g., Jeannerod, 2006; Pulvermüller & Fadiga, 2010), we might also expect evidence for motor resonance when people understand language about actions. In fact, there is strong evidence for cross-talk between the action and language systems, both at the neural and behavioural levels (see section 2.2.3 – 2.2.4). However, questions about necessity and direction of causality between the action and language systems remain unanswered. In the coming sections, I outline the current state of the field concerning embodied approaches to action language comprehension. First, in section 2.3.1, I explore to what extent language comprehension can be considered a form of action perception. In section 2.3.2, I sketch some current positions about the relation between language and action, and their predictions. In sections 2.3.3- 2.3.5, I assess to what extent these predictions are supported by current data.

2.3.1. *Language comprehension as action perception*

In recent years, the rise in research on action perception, in particular mirror neurons, has led to researchers viewing language as action – and, language comprehension as a form of action perception (Pickering & Garrod, 2013). This apparent parallel between action perception and language comprehension has prompted researchers to investigate the role of motor resonance in language. However, Fischer and Zwaan (2008) warn against straightforwardly viewing language comprehension as a form of action perception, for three reasons.

First, they argue that whereas the speed of action observation is determined by the speed at which the action is performed, the speed of language processing is determined

instead by the speech rate of the speaker or reading rate of the reader. However, some evidence suggests that in spoken language, speech rate appears to correlate with the speed of the described event (Shintel, Nusbaum, & Okrent, 2006); and in written language people tend to be faster to respond to sentences describing short distances, compared with long distances (Matlock, 2004). Similarly, in a recent study using eye-tracking, participants listened to sentences such as *the student ran/ staggered along the trail to the picnic basket* while observing a visual array featuring the actor, the path, and the actor's target location (Lindsay, Scheepers, & Kamide, 2013). Participants looked more often and for longer periods at the path in sentences in which the verb implies slow movement along the path (e.g., *stagger*), compared with sentences in which the verb implies fast movement along the path (e.g., *run*). These results suggest that, contrary to Zwaan and Fischer's claim, the speed at which language is processed *is* in fact affected by the implied speed or duration of the action being described. However, these results only describe a very general effect of processing speed, comparing a "fast" condition against a "slow" condition. When observing an action on the other hand, the rate at which that action is understood is necessarily *proportionate* to the speed at which the action is performed. Future research should therefore aim to establish whether there is evidence of a similar proportional relationship between language processing and the speed described in a sentence.

Fischer and Zwaan's second point is that in sentence processing, unlike action perception, understanding occurs only when a uniqueness point is reached, and it is clear what particular action is being described. For example, in the sentence *John is opening the can*, it is only once the comprehender hears the word *can* that she can activate a relevant motor representation: the verb *opening* could refer to many actions, all of which require entail differential involvement of the motor system (e.g., opening a door, opening a museum). Therefore, motor resonance should arise only once the uniqueness point is reached. This prediction has received empirical support from studies showing that motor resonance does appear to occur only when it becomes clear which particular action is being

described (e.g., clockwise or anticlockwise rotation; Taylor, Lev-Ari, & Zwaan, 2008). However, this uniqueness point also seems to be present in action perception. For example, when I observe John reaching towards a desk, I may not know until quite late into his reaching, whether he is planning to pick up his coffee cup (i.e., execute a power grip) or his pencil sharpener (i.e., execute a precision grip). Given that motor resonance is sensitive to the distinction between power and precision grip (e.g., Rizzolatti & Craighero, 2004; see also section 2.2.3), this means that in action perception, as in action language, motor resonance – or at least, motor resonance specific to the action being performed – occurs at a critical point, once there is sufficient evidence to predict one action over another. Of course, in many cases, the context of the action will make it clear whether John is intending to pick up his coffee cup or his pencil sharpener; but outside of laboratory conditions, this is also the case for language. It would usually be quite clear from the context of the conversation whether the speaker is describing John opening the can, or the museum, before the disambiguating word (*can/ museum*) has been reached. In this situation, we would expect motor resonance to occur before the disambiguating word (see Kaschak & Borreggine, 2008 for relevant discussion).

Fischer and Zwaan's third point is that unlike action observation, in which an observer necessarily sees the action in all (or most of) its entirety, language is often underspecified. For example, in the sentence *He turned the page*, it is not clear who is doing the turning, and what sort of page is being turned (e.g., book page, calendar page). However, I argue that this is a feature of language use in experimental settings (i.e., isolated sentences presented without context). Research suggests that for longer stretches of language, comprehenders appear to build a situation model in which wider contextual information can be stored (e.g., Johnson-Laird, 1980; Zwaan & Radvansky, 1998). In typical language use, therefore, in which context is accumulated, it *would* be clear who was turning what sort of page. In this thesis, I will investigate the role of motor context (previous actions versus planned actions) in sentence comprehension. I therefore minimise the linguistic context

available to participants, by providing them with simple action sentences without the situational knowledge that would usually be present in naturalistic language comprehension. In this way, I hope to isolate the effects of motor activation on perspective taking in sentence comprehension. Once this has been established, future work should, of course, look at how the role of motor activation develops in more naturalistic language settings where linguistic context is available for the comprehender (see section 8.3.4 for further discussion of the role of motor resonance in naturalistic language comprehension).

In summary, Fischer and Zwaan (2008) provide three reasons for arguing that language comprehension involves a different set of constraints on motor resonance, compared with action perception. However, on closer inspection, it seems that the constraints described by Fischer and Zwaan are in fact more general constraints on action perception, and not specific to language comprehension. It is clear, however, that there is something different about language compared with action observation: namely, that language involves both communicative and referential motor resonance (see section 1.1). Throughout the thesis, therefore, I will consider language comprehension as a form of action perception, albeit a slightly special type of perception, which involves not one but two levels of motor resonance. In the next section, I outline various approaches to referential meaning in language, and the claims that these accounts make for the link between action and language.

2.3.2. Types of embodiment theory

In a recent review article, Meteyard, Cuadrado, Bahrami and Vigliocco (2012) placed theories of semantics on a continuum from unembodied accounts at one end, to strongly embodied accounts at the other. Unembodied approaches propose that semantics is totally amodal, and completely independent from the sensorimotor system (e.g., Latent Semantic Analysis; Landauer & Dumais, 1997); such accounts are strongly challenged by evidence for cross-talk between the language and motor systems (sections 2.3.4 and 2.3.5). Next on the continuum, secondary embodiment approaches are typically be framed as non-

embodied in the literature, because they propose that semantic systems are amodal in nature, and they reject the notion of dependence between the sensorimotor and language systems (e.g., Mahon & Caramazza, 2008). However, secondary embodiment accounts do allow for associated activation between the language and the action systems. For example, Mahon and Caramazza (2008) accept that spreading activation might occur between the two systems (i.e., activation of an abstract, amodal concept can trigger cascading activation in the motor system). What secondary embodied accounts deny is that this motor activation plays any functional role in semantic processing. In other words, secondary embodied accounts predict that there may be temporal or spatial overlap between activation of the language and motor systems, but there should not be any major disruption to one system when other system is impaired (Meteyard et al., 2012).

In contrast, weak embodiment accounts propose that sensory and motor information contributes to (but does not wholly constitute) semantics (e.g., Pulvermüller, 1999; Vigliocco, Vinson, Lewis, & Garrett, 2004). These theories also predict some temporal and spatial overlap between activation in the motor and language systems, and, in addition, claim that motor activation does play a representational role in meaning; it is more than an epiphenomenon (cf. Mahon & Caramazza, 2008). According to Meteyard and colleagues, weak embodiment accounts therefore predict that disruption in one system would result in disruption in the other system, but they do not claim that motor activation is *sufficient* for sentence comprehension to occur. For example, we might construe Taylor and Zwaan's (2013, 2009) recent proposal of fault-tolerant comprehension as weakly embodied: in this view, language comprehension does not fail in the absence of motor simulation, but rather, degrades gracefully, perhaps resulting in a slightly less rich representation of the described action. The proposal is attractive because it explains how motor resonance might play a role in language comprehension, while still allowing for comprehension to occur in the absence of appropriate motor activation (i.e., it does not commit us to the view that participants fail to understand linguistic stimuli in non-activation conditions).

Weak embodiment accounts also allow for a degree of abstraction. For example, Simmons and Barsalou (2003) propose that information from different modalities is bound together to constitute semantics – as Shapiro (2010) points out, this binding process requires abstraction. This hypothesised binding process also has implications for the localisation of action language semantics, since the “bound” representations are no longer modality specific, and should therefore be located *adjacent to* modality specific areas (Meteyard et al., 2012). Therefore, weak embodiment accounts predict activation in areas adjacent to the primary motor cortex, such as the pre-motor cortex.

Finally, Meteyard et al. characterise strong embodiment approaches by a complete dependence between meaning and sensorimotor activity (e.g., Glenberg & Kaschak, 2002), in which simulation is both necessary and sufficient for language comprehension: the activation of the motor system during action language comprehension *is* comprehension. Unlike weak embodiment accounts, strong embodiment accounts therefore predict that activation will occur in the same regions during action execution and action language understanding (i.e., the primary motor cortex). Interestingly, Meteyard et al. note that theories of strong embodiment tend to be concerned with sentence and narrative processing, which intuitively lend themselves to the idea of simulation, through situation models. For example, the Indexical Hypothesis (Glenberg & Kaschak, 2002; Glenberg & Robertson, 2000) proposes that comprehenders understand language by combining object affordances (the potential for interaction that an object presents; Ellis & Tucker, 2000) into an action plan; by simulating performing these actions, comprehenders understand the meaning of the action described in the sentence. However, note that a series of fMRI studies have suggested that activation of the motor system is stronger following isolated verbs (e.g., *grab*) than following third-person sentences (e.g., *The fruit cake was the last one so Claire grabbed it*; Raposo, Moss, Stamatakis, & Tyler, 2009), contradicting the suggestion that simulation may be more likely to occur in sentences than in single words. In Chapter 3, I discuss in more detail the possible role of situation models in action language.

One of the difficulties with placing accounts on a continuum such as that of Meteyard et al. (2012), is that accounts vary along several dimensions (see also Kiefer & Pulvermüller, 2012). Therefore, it is that it is not always clear where an account should sit. In particular, the distinction between strong and weak embodiment is not clear-cut. For example, Barsalou's Perceptual Symbols Systems (1999) could be classified as either strongly or weakly embodied, depending on which dimension (abstraction, or dependence) is weighted most heavily: the claim that language comprehension does not involve complete reenactments of previous motor activity, but only some attention-driven subsets of that activation, implies a degree of abstraction in line with weak embodiment. On the other hand, the claim that simulation *constitutes* understanding is strongly embodied. Compounding this problem is the fact that embodiment has typically been framed as an alternative to unembodied accounts of language. Therefore, embodied theories have rarely specified whether they are arguing the case for weak or strong embodiment. Many studies accept activation in the motor and in the pre-motor cortex as equally supportive of embodiment (either weak or strong), when in fact, as Meteyard et al. have pointed out, different types of embodied theory predict different loci of activation. In the following section, we discuss evidence for shared neural resources between the language and motor systems.

2.3.3. Neural evidence for shared resources

The neural exploitation hypothesis states that brain mechanisms adapt to serve new functions in addition to their original roles; in the case of motor resonance and language, motor systems have adapted to perform a role in language comprehension as well as action execution and planning (Gallese & Lakoff, 2005; Gallese, 2009; see also Anderson, 2010). In the following sections, we outline evidence that language and action activate common neural substrates. The implication of such evidence is that one system exploits these shared resources in order to aid its own processing – for example, the language system might exploit motor areas by using action simulation to improve comprehension.

Language areas activated by motor tasks

There is increasing evidence that areas of the brain typically recruited during language processing and production are also recruited when performing motor tasks. Broca's area is usually defined as comprising the pars opercularis and pars triangularis in the left inferior frontal gyrus, or Brodmann's areas 44 and 45 (e.g., Dronkers, Plaisant, Iba-Zizen, & Cabanis, 2007). This area has traditionally been ascribed an important role in sentence processing (e.g., Dapretto & Bookheimer, 1999; Friederici, 2002). However, recent work has also implicated this area in motor representations. For example, research using PET showed activation during observation and mental imagery of grasping movements (Grafton et al., 1996), and studies using fMRI found activation while executing and imagining object manipulation (Binkofski et al., 1999; Gerardin et al., 2000).

Broca's area is often considered a homologue of area F5 in monkeys, where the mirror system is thought to be located (e.g., Rizzolatti & Craighero, 2004; Rizzolatti & Sinigaglia, 2010). Like mirror neurons in monkeys, there is evidence that action-based activation in human Broca's area only occurs following meaningful hand to object interactions, rather than hand movements in general (Decety & Grèzes, 1999; Johnson-Frey et al., 2003). These similarities between Broca's area in humans and area F5 in monkeys have led some researchers to ascribe Broca's area an important role in the evolution of speech, with a pre-linguistic grammar of gestures later evolving into oro-facial communication and vocalisation (Arbib, 2005; Rizzolatti & Arbib, 1998; see Corballis, 2013, for a recent review).

On the other hand, we must be wary of using evidence that language areas are activated during action as support for the claim that motor resonance is involved in language processing. The results highlighted above may mean no more than that areas traditionally associated with one function are in fact multifunctional. For example, Broca's area is also activated during musical processing (see Fadiga et al., 2009 for a review) but this is not usually taken as evidence that the "musical system" is involved in language comprehension.

Motor areas activated by language tasks

Most arguments in favour of motor resonance in language processing are therefore based on evidence suggesting shared resources in the other direction (i.e., that language processing results in activation of motor areas). For example, in a PET study, Vigliocco et al., (2006) found that activation in the premotor and primary motor cortex increased when participants listened to motor words (both verbs and nouns) compared with sensory words (both nouns and verbs). Oliveri et al., (2004) also crossed meaning with grammatical class, using TMS to measure hand MEP responses while participants performed morphological transformations on action- or non-action-related words. Action words mostly (but not entirely) referred to hand actions; hand MEPs increased for action compared with non-action words, suggesting that participants had accessed some of the motor properties of the words.

In further support of the hypothesis that language comprehension involves activating motor systems, fMRI research has shown increased activation in motor and premotor areas for action words compared with object words (Kable, Kan, Wilson, Thompson-Schill, & Chatterjee, 2005; Kable, Lease-Spellmeyer, & Chatterjee, 2002), and action words compared with abstract words (Noppeney, Josephs, Kiebel, Friston, & Price, 2005). Further work suggests activity in action-related regions for tool words compared with non-tool words (Martin & Chao, 2001; Martin, Wiggs, Ungerleider, & Haxby, 1996), and for manipulable versus non-manipulable objects (Saccuman et al., 2006).

Of course, if we argue that language activation during motor tasks need not imply any interaction between the two systems, only that certain areas of the brain are multifunctional; then we can – and perhaps, should – also deny that motor activation during language processing implies any interaction, for similar reasons. However, the suggestion that language comprehension involves simulations of the described actions is given additional credence by research suggesting that the motor activation occurring during language processing is somatotopic.

Somatotopic activation

A key feature of both weak and strong embodied approaches to language is that both types of account predict a degree of specificity in motor activation, corresponding to the action being described (e.g., Pulvermüller, Shtyrov, & Ilmoniemi, 2005). The motor strip is divided into precise areas corresponding to particular body parts (e.g., Pennfield & Rasmussen, 1950; but see Schieber, 2001, for constraints on this somatotopic organisation). We saw in section 2.2 that observing an action leads to specific activation of the body parts involved in performing that action (e.g., Buccino et al., 2001). Following on from these observations, embodied approaches to language make the testable prediction that the sensorimotor systems will be recruited during language processing in a precise, somatotopic manner (Chatterjee, 2010).

Using fMRI, Hauk, Johnsrude and Pulvermüller (2004) found that listening to leg-related words (e.g., *kick*) produced activation in areas of the motor strip responsible for planning and executing leg movements; arm-related words (e.g., *pick*) produced activation in areas associated with arm movements, and mouth-related words (e.g., *lick*) produced activation in areas associated with mouth movements. Similar fMRI findings have been described for spoken sentences relating to hand, foot or mouth movements, compared with abstract sentences (Tettamanti et al., 2005), and for visually presented action sentences compared with metaphorical sentences (Aziz-Zadeh, Wilson, Rizzolatti, & Iacoboni, 2006). In a study with TMS, Buccino et al., (2005) found that hand and foot MEPs decreased when participants listened to hand- or foot-related sentences, respectively. All of these results appear to support the suggestion that the motor system is recruited during language comprehension, and that this recruitment is specific to the effector (arm, leg, mouth) implied by the described action.

However, these findings (and in particular, their use as evidence that action language comprehension is based on motor simulations) are not uncontroversial. As was the case with action perception and execution, the fact that fMRI and PET show activation in regions of

interest, rather than individual cells, means that we are unable to conclude whether the same cells (rather than different, action- or language-specific cells in the same region) are activated in action and in language processing. There is also evidence that the somatotopic organisation described above is in fact coarse grained, with overlap between effectors both within and between studies (Aziz-Zadeh & Damasio, 2008; Chatterjee, 2010; Fernandino & Iacoboni, 2010). Other researchers have failed to find evidence of somatotopy at all (Arévalo, Baldo, & Dronkers, 2012; Postle, McMahon, Ashton, Meredith, & de Zubicaray, 2008).

Another major limitation of neuroimaging techniques such as fMRI and PET is that they cannot tell us the time course of activation: does language processing precede motor activation, or does motor activation precede language processing? For example, in fMRI, the blood-oxygenation-level-dependent (BOLD) signal typically returns to baseline 20 seconds after stimulus presentation. This means that we do not know whether the observed activation reflects an early lexico-semantic stage of processing, or a later stage involving mental imagery (e.g., Tomasino, Werner, Weiss, & Fink, 2007). Fortunately, other techniques (MEG, TMS and EEG) operate at a high level of temporal resolution (see Hauk, Shtyrov, & Pulvermüller, 2008, for a review). For example, face- and leg- related words were presented while participants underwent MEG, showing evidence for somatotopic activation from around ~170 ms (i.e., an early processing stage typically associated with lexical processing; Pulvermüller, Shtyrov, et al., 2005). Similar evidence of early activation (below 250 ms) has also been found using EEG (Pulvermüller, Härle, & Hummel, 2001; Shtyrov, Hauk, & Pulvermüller, 2004).

Localising motor activation: The need for precision

It is sometimes argued that although functional neuroimaging can help us localise particular processes, competing theories do not often predict differences in localisation (e.g., Page, 2005). However, as we saw in section 2.3.2, strong embodiment predicts activation in

the primary motor cortex itself, whereas weak embodiment predicts activation in areas that are related to, but not the same as, those involved in executing the corresponding actions (e.g., premotor cortex¹). Can we use these differing predictions to distinguish between strong and weak embodiment accounts? In short, no. In addition to concerns over the specificity of somatotopic activation, there is also large variability in which parts of the motor system are activated in different studies. Researchers commonly report activation of the *motor circuits* (e.g., Tettamanti et al., 2005) or *motor system* (e.g., Marino, Gallese, Buccino, & Riggio, 2012; Pulvermüller, Hauk, Nikulin, & Ilmoniemi, 2005) during language processing, but the exact part of the motor system in which they find this activation varies considerably between studies. Some studies find activation in the primary motor cortex, as predicted by strong embodiment accounts (Hauk et al., 2004; Vigliocco et al., 2006), but others find activation in the premotor cortex (Aziz-Zadeh et al., 2006; Boulenger, Hauk, & Pulvermüller, 2009), pre-supplementary motor area (Lyons et al., 2010), or inferior parietal lobule (Péran et al., 2010), as predicted by weak embodiment accounts.

In spite of this variable localisation, researchers have tended to assume that the same underlying processes are at work (i.e., motor simulation), when, in fact, this need not be the case (Chatterjee, 2010). For example, Willems, Toni, Hagoort and Casasanto (2010) found that a lexical decision task produced activation in premotor areas, while a mental imagery task on the same words produced activation in the primary motor cortex, and non-overlapping parts of the premotor cortex. From these differences in localisation, Willems et al. concluded that *different* cognitive processes were taking place. This example is striking because much of the debate over the role of the motor system in action language comprehension centres on the question of whether the observed effects are simply the result of explicit motor imagery (e.g., Tomasino et al., 2007).

In this section, I have outlined evidence for shared neural resources between the language and motor systems. I have also outlined some concerns regarding this evidence,

¹ Assuming that neuroscience has correctly mapped out the functions of these areas

including a lack of temporal resolution, and considerable variable in spatial localisation. Despite such concerns, cognitive psychology papers incorporating neuroimaging results are typically seen as more convincing than those without (McCabe & Castel, 2008; Weisberg, Keil, Goodstein, Rawson, & Gray, 2008), and as a result, the studies outlined above have been hugely influential in garnering support for embodied approaches to language. However, embodied accounts do not only claim shared resources between the language and motor systems; they also argue for various degrees of dependency relation between the two systems. In other words, embodied approaches predict interaction between the two systems. To find evidence of an interaction, we must turn to behavioural research.

2.3.4. Effects of language on action

The neural evidence outlined above seems to discount what Meteyard et al. (2012) termed unembodied accounts, but it does not provide any evidence of interference or facilitation between the action and language systems that might help us distinguish between secondary, weak, and strong embodiment approaches. In fact, there is substantial evidence that language affects the speed, accuracy, or kinematics of action execution, in much the same way as action perception does. For example, we saw in section 2.2.2 that performing an action increased perceptual accuracy for congruent actions (Bidet-Ildei, Chauvin, & Coello, 2010). Bidet-Ildei, Sparrow and Coello (2011) extended this work by displaying an action word for 500 ms prior to observing actions in a point light display. When the movement stimulus was preceded by a congruent action verb, participants were faster to detect the presence of a human rather than random motion in the stimulus.

Similarly, recall from the action-perception literature that Hamilton et al. (2004) found that participants tended to judge a box lifted by someone else as heavier when they themselves were lifting a light box, and to judge a box lifted by someone else as lighter when they themselves were lifting a heavy box. Hamilton et al. explained these findings by appealing to the MOSAIC model of action control (Haruno, Wolpert, & Kawato, 2001;

Wolpert & Kawato, 1998) whereby a module for a specific task (e.g., lifting a heavy box) can be rendered temporarily unavailable to the perceptual system when it is being used to execute an action (see also discussion of blindness to response-compatible stimuli in section 2.2.3). Scorolli, Borghi and Glenberg (2009) extended this box-lifting paradigm to study the effect of language on action. They had participants lift identical-looking heavy or light boxes after listening to sentences describing lifting a light object (e.g., *pill*) or a heavy object (e.g., *chest*). The logic of the experiment is that, if understanding action language involves activating the motor system responsible for executing that action, then understanding a sentence like *move the pillow from the ground to the table* should render the module for lifting light objects temporarily unavailable, resulting in a delay for participants lifting a light object, but not a heavy object, following this sentence. Indeed, Scorolli et al. found an interference effect in lift delay: participants were slower to lift the box having grasped it when the weight implied in the sentence was congruent with the actual weight of box. The fact that language shows similar effects as action observation suggests that language and action (or at least, action perception) share common mental representations.

Action-sentence compatibility effects

One of the most well-known phenomena in this area of research is the Action-Sentence Compatibility Effect (ACE; Glenberg & Kaschak, 2002). In this paradigm, participants read a sentence describing an action in a particular direction (e.g., towards or away from the body), and make a sensibility judgement on that sentence by performing an action that is congruent or incongruent with the direction of the sentence (e.g., press button nearer or further from their body). A facilitation effect is typically found when the sentence and response direction are congruent (e.g., pressing a button further from the body in response to the sentence *Close the drawer*). As well as imperatives such as *Close the drawer*, Glenberg and Kaschak (2002) also found the ACE for concrete transfer sentences (e.g., *you delivered the pizza to John*), and abstract transfer sentences (e.g., *Liz told you the story*). The

ACE has since been replicated in third-person sentences such as *Lea delivers the pizza to Lea*, (Gianelli, Farnè, Salemme, Jeannerod, & Roy, 2011), although only when spatial avatars are provided for the agent (see Chapter 3 for further discussion). The ACE has also been found using auditory sentences, in a go-no go paradigm (Borreggine & Kaschak, 2006), and in American sign language (Secora & Emmorey, 2013). Aravena et al., (2010) recently provided evidence for a neural signature of the ACE using open and closed handshapes rather than direction of movement.

Furthermore, Glenberg et al. (2008) extended work on the ACE by applying TMS to measure MEPs while participants responded to Italian sentences describing no transfer, concrete transfer, or abstract transfer. For technical reasons, MEPs could not be measured for towards and away movements, and so instead the authors measured grasping-related activation both during and after sentence processing. Findings showed increased activation when TMS was delivered at the end of the verb compared with the end of the sentence; and increased activation for transfer versus non-transfer sentences. There was no difference between concrete and abstract transfer sentences. These results suggest that the ACE occurs equally for concrete and abstract transfer sentences, and that motor resonance occurs during, rather than after sentence processing. Results from Zwaan and Taylor (2006) also suggest that motor resonance is a short-lived, localised phenomenon: participants turning a knob in a self-paced reading paradigm were faster to turn the knob when the verb matched the direction of turning (clockwise, anti-clockwise); but the effect was restricted to the verb in the sentence and did not spill over onto later words.

Timing and effect direction

In section 2.2.2, I discussed the fact that the direction of effects found in automatic imitation was inconsistent, and the temporal dynamics might play an important role in determining the direction: when participants observe an action and subsequently perform an action, results tend to show a facilitation effect; but when participants observe and execute

an action simultaneously, interference effects have been found (e.g., Kilner et al., 2003). The results in action language comprehension seem even less consistent. For example, when Buccino et al. (2005) stimulated arm or leg areas using MEPs as participants listened to arm- or leg-related sentences, they found reduced MEPs in the congruent condition. In a follow up behavioural study using semantic categorisation, they found slower RTs when participants responded using an effector that was congruent with the action described in the sentence. On the other hand, Pulvermüller, Hauk, et al., (2005) found *increased* MEPs in the congruent condition when they applied TMS to leg- or arm-areas. Similarly, Kaschak et al. (2005) found that participants were slower to make sensibility judgements on sentences while watching motion in a congruent direction; but a study using the same linguistic materials found that when the implied motion was presented auditorily, instead of visually, participants were *faster* to make sensibility judgements in the congruent condition (Kaschak, Zwaan, Aveyard, & Yaxley, 2006; see section 2.4.1 for a review of effects in behavioural research).

Could temporal dynamics help explain the inconsistency in action language research as well as in automatic imitation? The studies by Glenberg et al. (2008), and Zwaan and Taylor (2006) outlined above suggest that the ACE occurs at specific points in time rather than throughout the entire sentence. The direction of response might therefore depend on the point in time at which a response is made. For example, Boulenger et al. (2006) found evidence that action verbs affect movement kinematics, and that this effect reverses depending on the timing constraints. A stimulus was presented on screen just after participants began a grasping movement in a go/ no go task; if the stimulus presented was a word, participants had to complete the grasping action, and if it was not a word, participants had to interrupt the grasping action. Results showed that when action verbs (but not concrete nouns) were presented, participants' hand acceleration decreased, implying an interference effect of language on action. The effect was stronger for hand than for non-hand related action verbs (but note that the effect was still present, albeit weaker, for non-hand action

verbs, implying some degree of abstraction over the effector used). Boulenger et al. (2008) used subliminal display of words to rule out conscious imagery of the described action as the cause. However, when Boulenger et al. (2006) presented the word/ non-word stimulus *before* participants initiated their grasping action, hand acceleration *increased* rather than decreased, suggesting a facilitation effect. Willems and Hagoort (2007) note that Boulenger et al.'s results might help explain the discrepancy between the lowered MEPs in Buccino et al. (2005)'s TMS study, and the increased MEPs in other studies (e.g., Oliveri et al., 2004; Pulvermüller, Hauk, Nikulin, & Ilmoniemi, 2005). In Buccino et al.'s study, TMS was applied during sentence presentation, so that language processing was not yet complete (as in Boulenger's Expt. 1); in the studies by Pulvermüller et al. and Oliveri et al., TMS was applied once language processing was complete (as in Boulenger's Expt. 2).

Affordances

The term *affordances* originally dates from Gibson (1977), where it refers to the action possibilities afforded to a particular organism by objects or the environment (e.g., a tree might afford climbing to a squirrel, but not to a whale). More recently, the term has been used to describe automatically elicited action possibilities, which are related to specific *components* of action, for example grasping with a power versus a precision grip (Symes et al., 2008; Tucker & Ellis, 2001, 2004). Over the past decade, researchers have become increasingly aware that language can affect how people carry out particular actions, by tapping into these action-relevant properties (affordances) such as orientation or size. For example, Creem and Proffitt (2001) found that participants asked to recall one word from a pair of semantically related words (e.g. *celery – pear*), were less likely to grasp a tool (e.g., a spatula) in a way that would afford normal tool use, compared with participants who completed a visuo-spatial or no concurrent task. However, note that the word pairs used by Creem and Proffitt were drawn from a range of categories (e.g., food, furniture, animals) and were not controlled for semantic relatedness to tool use. The results therefore suggest a

possible effect of language *in general* on action, but not the semantic specific interference that embodied theories (strong or weak) would predict. In fact, it may be the case that the semantic task was simply more challenging than the visuo-spatial task.

Evidence for a more specific effect of language on action comes from a recent study in which participants performed a lexical decision task where the critical words denoted objects that are typically used by being brought towards the body (e.g., *cup*), or away from the body (e.g., *key*). Participants were faster to respond when the direction of response (towards or away from body) matched the direction in which the object was typically moved (Rueschemeyer, van Rooij, Lindemann, Willems, & Bekkering, 2010). These results suggest that semantic information about an object can activate motor programmes connected with the typical affordances for that object (see also Lindemann, Stenneken, van Schie, & Bekkering, 2006).

Interactions between language and affordances also occur when semantic processing is irrelevant to the task. For example, participants were more likely to use a power grip to pick up a wooden block when they had just read a word referring to an object typically manipulated with power grip (e.g., *apple*); and more likely to use a precision grip to pick up an identical block when they had just read a word referring to an object typically manipulated using a precision grip (e.g., *grape*; Glover, Rosenbaum, Graham, & Dixon, 2004). Similarly, Gentilucci and Gangitano (1998) had participants reach for and grasp a rod on which the Italian word for *long* or *short* was printed. The kinematics of participants' reaching movements tended to approach the rod as near or far, according to whether it featured the word *long* or *short*, even though word reading was irrelevant to the task (see also Gentilucci, Benuzzi, Bertolani, Daprati, & Gangitano, 2000). Glover and Dixon (2002) found that this affect occurred early on but dissipated as participant's hands approached the stimulus, implying that the effect was restricted to the early planning stages of movement.

We saw in section 2.3.2 that affordances play an important role in Glenberg's Indexical Hypothesis (Glenberg & Kaschak, 2002; Glenberg & Robertson, 2000; Kaschak &

Glenberg, 2000), which claims that comprehenders combine affordances derived from language into potential action plans, which are used to simulate the actions described in the sentence. In support of the Indexical Hypothesis, Borghi (2004) had participants read a sentence (e.g., *the woman shares the orange*), followed by a noun (e.g., *slice* or *pulp*). The task was to decide whether or not the noun referred to part of the object mentioned in the sentence. When a noun was easier to derive an affordance from (e.g., it is easier to derive the act of sharing from *slice* than from *pulp*), the subsequent word was processed more quickly, even though both nouns were semantically related to the critical word in the sentence. These results support the Indexical Hypothesis by suggesting that we understand sentences through processing the affordances of the objects being referred to, rather than through associative relations between words.

One relevant observation about affordances, made by Borghi and Riggio (2009), is that they can remain constant across different contexts (*stable* affordances; e.g., object size), or they can change according to the context (*temporary* affordances; e.g., object orientation). A canonical affordance is a special kind of affordance, which captures the way in which we typically *use* (rather than, for example, move or clean) a given object. For instance, a calculator might be passed to a friend the right way up or upside down (temporary affordance), but our *use* of a calculator is typically restricted to cases where it is right way up (canonical affordance). Crucially then, canonical affordances are based on our previous experience. Borghi argues that action perception recruits temporary affordances, as an object is presented to us at that point in time. However, in language processing, we construct a motor prototype, based on canonical affordances (Borghi & Riggio, 2009; Borghi, 2013). Thus, in the Indexical Hypothesis, comprehenders build potential action plans based on the way that they have typically interacted with objects in the past. Note that this recruitment of canonical affordances in language entails a degree of abstraction over accumulated sensorimotor experience, just as Barsalou's perceptual symbols entail an abstraction over the various instances of single modality input to create a coherent, multimodal representation.

2.3.5. Effects of action on language

Evidence outlined in the previous section suggests that language does affect motor responses, in particular by affecting the way we process a describe object's affordances. However, what embodied accounts of language processing (both weak and strong) would really like to show is that the relationship is two-directional – in other words, an effect of motor resonance on language processing – since this would support the claim that language is at least partly dependent on the sensorimotor system. There are main two sources of evidence for a functional role of the motor system in language processing. First, language processing in patients with motor impairments; and second, language processing in healthy participants whose motor system has been temporarily activated as part of an experimental design. In the following sections, I briefly review these two literatures.

Patient studies

A feature of patient studies through the years has been a reliable double dissociation between the language and action systems (e.g., paralysis with intact language function, versus aphasia with intact motor function; Pulvermüller, 2005). This double dissociation seems to rule out strong embodied accounts that posit a total reliance of meaning on the relevant sensorimotor systems.

However, a range of studies examining specific action verb deficits in patients with reduced motor control, do suggest some language impairment, in support of weak embodiment theories. For example, Neiningen and Pulvermüller (2003) found that in a lexical decision task, patients with lesions in the right temporo-occipital cortex (associated with visual processing) showed a relative disadvantage in accuracy for nouns with strong visual associations; and patients with hemiparalysis, and lesions in the right frontal cortex (associated with motor processing), showed a relative disadvantage for verbs with strong motor associations. An impairment for verbs relative to nouns has also been observed in patients suffering from motor deficits due to motor neuron disease (Bak, O'Donovan,

Xuereb, Boniface, & Hodges, 2001), corticobasal degeneration and supranuclear palsy (Cotelli et al., 2006), Parkinson's disease (Cotelli et al., 2007; Péran, Rascol, Celsis, Nespoulous, & Dubois, 2003), and Huntington's disease (Péran, 2004). Importantly, Boulenger, Mechtouff, et al. (2008) found that the effect of masked priming on lexical decision for action verbs compared with concrete nouns was stronger in Parkinson's patients who were on, compared with off, dopamine treatment (known to improve motor cortex function; see Boulenger, Mechtouff et al. for discussion). However, unlike studies on healthy controls (Oliveri et al., 2004; Vigliocco et al., 2006), none of the above studies crossed grammatical class (verb, noun) with meaning (action, non-action). Therefore, their results cannot differentiate between a selective deficit for verb processing (caused by damage or degradation to the language system), and a selective deficit for action-related words (caused by damage or degradation of the motor system).

A more recent study (Péran et al., 2009) attempted to dissociate these two factors by using fMRI to measure activity as participants named images of objects, or produced verbs associated with the depicted object. The objects could be functionally manipulated (e.g., *key*) or could not be functionally manipulated (e.g., *bee*). Therefore, verbs produced in response to manipulable objects (e.g., *key – turn*), would be expected to have a higher actionable-content than verbs produced in response to non-manipulable objects (e.g. *bee – buzz*). Behavioural results showed that reaction times for verb generation were slower for manipulable than for non-manipulable objects, suggesting an additional impairment in producing verbs with an actionable semantic content, over and above a general deficit for verbs. Importantly, the authors found a significant positive correlation between the severity of motor deficit (as revealed by pre-tests) and the amount of activation in motor and pre-motor cortex during verb generation.

Mahon and Caramazza (2008) suggest that patient studies such as those of Boulenger et al. (2008), and Neiningen and Pulvermuller (2003), provide some of the strongest evidence against the argument that activation of the motor system is irrelevant to

lexical processing. However, note that both of these studies used lexical decision, a task which does not usually require deep semantic processing, and indeed for which other studies on healthy participants have failed to find embodied effects (Sato, Mengarelli, Riggio, Gallese, & Buccino, 2008). A further point to bear in mind when considering patient studies, is that right sided lesions resulting in impairment of action language processing or semantic knowledge are extremely rare (Kemmerer, Rudrauf, Manzel, & Tranel, 2012). In other words, even if damage to the motor system does result in action-language processing deficits, this only seems to occur when the damage occurs in the language-dominant left hemisphere, implying that degradation of language-related structures might also be involved in such deficits. Finally, note also that the majority of patient studies investigate deficits in verb generation or production, compared with embodiment research on healthy participants, which looks almost exclusively at language comprehension. Overall therefore, patient studies, while certainly suggestive of some amount of crosstalk between the language and motor systems, can provide only limited evidence for the routine involvement of motor resonance in typical language comprehension.

Motor activation primes language representations

Somewhat surprisingly, given that embodied approaches predict an effect of the motor system on language processing, research in this area is much sparser than research looking at the effect of language on action execution. Two notable studies have investigated the possibility of such an effect through the use of TMS. Willems, Labruna, D'Esposito, Ivry and Casasanto (2011) found that right-handed participants were faster to perform lexical decisions on manual action verbs than non-action verbs following stimulation of the left premotor hand area. Pulvermüller, Hauk, et al. (2005) used TMS to stimulate leg- and arm-related areas of the primary motor cortex, prior to presentation of leg- or arm-related action words. Participants used lip movements to respond on a go-no go lexical decision task: they were faster to respond to arm-related words than leg-related words following stimulation of

arm-areas, and faster to respond to leg-related words than arm-related words following stimulation of leg-areas. As with the results of Willems et al, these effects were only observed when TMS was applied to the language dominant left-hemisphere; no such effects were found applying TMS to arm- or leg-related areas in the right hemisphere, or during sham stimulation. The authors concluded that activation in the motor system facilitates lexical access to words that share a motor feature, such as effector, with this activation. In other words, motor activation primes visual word recognition, through shared representations between the motor and language systems.

Turning to behavioural research, the most well-known evidence for an effect of action on language is a study in which participants spent ~20 minutes moving beans towards or away from their body. In a subsequent sensibility judgement, participants were faster to respond to sentences implying the same direction of movement (towards, away) as they had performed in the training session (Glenberg, Sato, & Cattaneo, 2008). Other behavioural evidence is limited. For example, Helbig, Graf and Kiefer (2006) used a semantic priming paradigm, and found that participants were faster to name target pictures of objects when a previously presented picture showed an object that was used in a similar way as the target object (e.g., *nutcracker – pliers*), compared with when the prime object was used in a different way (e.g., *nutcracker – frying pan*). Helbig et al. interpreted these findings in terms of the role of action representations in object recognition, but, since the task was object naming, the results could equally be interpreted in terms of action representations and lexical access (i.e., between the motor and language systems). However, objects that are used in similar ways tend to share visual features (Rueschemeyer, Lindemann, van Rooij, van Dam, & Bekkering, 2010). Therefore, it is difficult to rule out a priming effect between perceptual form and language (e.g., Huettig & Altmann, 2007; Huettig & McQueen, 2011), rather than an effect of action on language.

Rueschemeyer et al. (2010) therefore had participants perform a lexical decision task on words depicting objects that could or could not be functionally manipulated (e.g.,

calculator versus *bookend*), while rotating a disk in the right hand. The disc rotation allowed Rueschemeyer and colleagues to activate participants' motor system while not activating perceptual features associated with the described objects. The distinction between functionally-manipulable, and non-functionally-manipulable words was motivated by a small body of literature suggesting that *functional* information about how an object is used, is activated more quickly and reliably than *volumetric* information about how an object is moved (Bub, Masson, & Cree, 2008; Buxbaum, Kyle, Tang, & Detre, 2006; Masson, Bub, & Newton-Taylor, 2008). Functional words may therefore involve the motor system to a larger extent than do volumetric words (see Péran et al., 2009 earlier in this section). Rueschemeyer et al. found that participants were faster and more accurate to respond to the functional words, compared with the volumetric words, suggesting that activation of the motor system facilitated recognition of those words in which motor resources played a more important role (see also Witt, Kemmerer, Linkenauer, & Culham, 2010).

Importantly, by using disc rotation to control for perceptual form similarities, Rueschemeyer et al. (2010) abstracted away from the particular type of action associated with their target action-words. Therefore, while their study suggests that generalised activation in the motor system may affect processing of action-related words, the results fail to demonstrate the level of specificity that would be required to support the proposal that comprehenders simulate the particular action being described. In a study preceding the recent upsurge in interest in the role of motor representations in language processing, Klatzky, Pellegrino, McCloskey and Doherty (1989) cued participants to make one of four different handshapes (poke/ pinch/ palm/ clench), before asking them to make a sensibility judgement on short phrases such as *squeeze a tomato*. Participants were faster to make these judgements when the handshape required to perform the action in sentence was congruent with the handshape prime, compared with when it was incongruent with the prime. This study suggests that activating the motor system for a particular action, may facilitate comprehension of congruent action language: in other words, action representations can

prime linguistic representations. In Chapter 7, I describe three studies that build on this work, to investigate a causal role of motor activation in language processing.

Necessity and causation

One advantage of conceiving the relationship between motor activation and language processing as a type of priming, is that it allows us to posit that the two systems involve shared representations, while relieving us of the burden of motor activation being either necessary, or sufficient, for action language comprehension. Therefore, cases in which there is no relevant action representation, but in which language is (presumably) understood regardless, no longer present a problem. Consider the case of structural priming: exposure to a particular syntactic form can prime the language user to repeat that form in production, or to interpret future sentences using the same structure (for a review, see Pickering & Ferreira, 2009). However, nowhere in the structural priming literature is it claimed that such priming is either necessary or sufficient for language processing to occur – rather, the more modest claim is made that in those cases where priming is possible, it facilitates a particular interpretation of language. In this thesis, I therefore argue that where there is shown to be an effect of action on language, the relationship between the language and motor systems represents a type of priming. In other words, action representations do not *cause* action language comprehension, but they facilitate a particular type of interpretation, one that is in line with recent motor system activation

It will be useful at this point to describe exactly what is meant by the terms *simulation*, *embodiment*, and *understanding*, under the above proposal. *Simulation* is used in this thesis to refer to the unconscious, automatic (in the sense of Heyes, 2010) activation of the same motor systems that would be used to perform a (linguistically described) action. In other words, it is used interchangeably with referential motor resonance. *Embodiment* is used to refer to the view that motor activation plays some role in the understanding of language. The importance of this role will depend on the type of embodiment theory (weak or strong,

see section 2.3.2); however, my proposal is that motor activation is neither necessary nor sufficient for understanding. *Understanding* therefore retains the amodal symbols and abstract operations assumed in most traditional psycholinguistic theories of semantics; it is these symbols and operations that allow understanding to occur in the absence of relevant motor activation. When motor activation is present during, or perhaps prior to language processing, this activation primes the comprehender to produce a particular type of simulation (e.g., slicing a tomato with a left hand rather than with a right hand). The way in which a particular sentence is understood might therefore vary with the type of motor activation available, in the same way that an ambiguous high/ low attachment sentence might be understood to mean different things depending on whether it was preceded by a high or low attachment prime. In both cases however, understanding still occurs in situations where no prime is available. See section 8.3.6 for further an evaluation of this proposal based on the results from Experiments 1-9.

2.4. Flexibility in embodied representations

In sections 2.3.4 and 2.3.5, I outlined some of the considerable evidence for cross-talk between the language and action systems. Embodied approaches to language argue that this cross-talk reflects the fact that semantic representations of action language are at least partly constituted by action simulations in the motor system (e.g., Glenberg & Gallese, 2012; Taylor & Zwaan, 2009). Some researchers have argued that the link between language and action is immediate and automatic (Pulvermüller, 2005); this claim is supported by evidence that the link occurs early on in processing (Hauk & Pulvermüller, 2004; Pulvermüller, Härle, & Hummel, 2000), and remains even when processing the action words is irrelevant to the task (Pulvermüller, Shtyrov, & Ilmoniemi, 2005; Shtyrov, Hauk, & Pulvermüller, 2004; but see Hoedemaker & Gordon, 2013). However, we also saw in section 2.3.3 that there is potentially worrying inconsistency in the localisation of activation in neuroimaging research into action language processing (see Tomasino & Rumiati, 2013, for a review). Behavioural

research in embodied approaches to action language yields similarly inconsistent results, with different studies finding results in different directions (facilitation versus interference), or no effects at all (see Table 2-1 for a summary of findings). In particular, effects seem to vary depending on the linguistic context of the stimuli (first-person *versus* third-person; present tense *versus* past tense), and the strategic context of the task (deep processing *versus* shallow processing; blocked trials *versus* randomised trials). This variability, and its implications for the automaticity of embodied semantic representations, is discussed in more detail below.

2.4.1. Sensitivity to context

Researchers have attempted to explain inconsistencies in the behavioural data in terms of temporal dynamics (e.g., Borreggine & Kaschak, 2006; Kaschak & Borreggine, 2008), or the level of semantic processing required by the task (e.g., Sato, Mengarelli, Riggio, Gallese, & Buccino, 2008). A consultation of Table 2-1 suffices to show that the latter explanation cannot be the whole story – for example, whereas Sato et al. found no embodiment effects in a lexical decision task requiring relatively shallow semantic processing, other researchers did find behavioural effects in lexical decision (e.g., Boulenger et al., 2006; Rueschemeyer, Pfeiffer, & Bekkering, 2010; van Dam, Rueschemeyer, Lindemann, & Bekkering, 2010). An explanation in terms of timing of stimuli alone is similarly insufficient: Borreggine and Kaschak (2006) found a facilitative ACE when the response direction cue was presented at sentence onset; but Diefenbach, Rieger, Massen and Prinz (2013), using the same timings, observed an interference effect. Diefenbach and colleagues explain this discrepancy by pointing to the fact that although their response direction cue appeared on screen at the same time as that of Borreggine and Kaschak, it remained onscreen for longer, meaning that some participants may not have immediately paid it attention. The discrepancy between the results of Diefenbach et al. and those of

Borreggine and Kaschak provides a particularly striking example of the fact that effect direction appears to depend on theoretically irrelevant differences between studies.

The role of experimental power

One possible explanation for the inconsistency in findings, in particular where there are no consistent differences in experimental parameters such as timing or task, is variation in the amount of experimental power between different studies. The majority of studies in Table 2-1 did not report their results in sufficient detail to allow power analyses to be conducted. I therefore calculated the number of observations per cell in the experimental design in order to gain some comparable measure of power across studies. We can see that some inconsistencies might indeed be explained by differences in power. For example, Borghi and Scorolli (2009) found a null effect when randomising hand- and foot-related stimuli, and a significant congruency advantage when these stimuli were blocked. The authors attribute the null result to the randomisation procedure, but closer analysis reveals that there were only half as many observations in the randomised version as in the blocked version of the experiment. Clearly then, we cannot rule out that the null result was due to lower experimental power.

On the other hand, several other papers report two or more experiments showing differences in effect direction, but keep the number of observations per condition constant between studies (e.g., Bergen & Wheeler, 2010; Borreggine & Kaschak, 2006; Boulenger, 2006). In addition, although experimental power might explain inconsistencies in whether or not a result is significant, it seems more difficult to invoke experimental power to explain significant effects in opposite directions, as are seen throughout the embodiment literature. Therefore, although differences in experimental power should be checked, doing so reveals no clear pattern that might explain the inconsistency in embodiment findings. I therefore conclude that at least some of this inconsistency must be attributed to differences in

experimental parameters. Identifying what these parameters are is an important next step in embodiment research.

The role of strategy

Recently, Papeo, Rumiati, Cecchetto and Tomasino (2012) demonstrated that the motor system's response to words depends not only on the meaning of that word, but also on the type of strategy used in a previous, apparently unrelated, task. Participants underwent an fMRI scan while reading Italian verbs describing manual actions (e.g., *stir*) or psychological states (e.g., *adore*). Previous research has shown that manual action verbs elicit more activation in motor regions of the brain, compared with state or psychological verbs (e.g., Rueschemeyer, van Rooij, Lindemann, Willems, & Bekkering, 2010). However, before reading the verbs in Papeo et al.'s study, participants completed a mental rotation task relying on either a motor or a visual strategy (Wraga, Thompson, Alpert, & Kosslyn, 2003). Results showed that the previous task context (motor-based or visual-based strategy) affected activation in response to the verbs. Notably, state verbs following the motor-based mental rotation task showed increased activation in motor areas compared with state verbs following the visual-based rotation task. The authors suggest that these results have important implications for fMRI studies in which blocked designs mean that participants are first asked to perform an action execution task in order to "localise" motor areas, and then to perform a language based task.

It is important to note however, that such results, coupled with the inconsistency in findings outlined above, do not necessarily challenge the assertion that the link between language and action is automatic. Heyes (2011, p. 470) defines automaticity in relation to automatic imitation (i.e., visuomotor priming; see section 2.2.2) as meaning that the phenomenon occurs "independent of the actors' intentions". This definition of automaticity is compatible with the observed inconsistency in embodied cognition results, provided that the effect of context acts as a form of automatic priming that is not under the actor's strategic

control. What these results do mean, however, is that researchers in this field need to systematically explore *in which contexts* language appears to interact in particular ways with the motor system, and to incorporate these details into their theories.

The role of linguistic context

For example, linguistic context appears to affect how the language and motor systems interact with one another, although the exact manner in which this occurs is unclear. Some researchers have found significant interactions between language and action using third-person sentences such as *Dave removed the screw from the wall* (Taylor et al., 2008; Taylor & Zwaan, 2008; Zwaan & Taylor, 2006), but other researchers have found that third-person language appears to elicit different representations compared with first- and second-person sentences (Brunyé, Ditman, Mahoney, Augustyn, & Taylor, 2009; Brunyé, Ditman, Mahoney, & Taylor, 2011). I discuss the issue of self-referential versus third-person language in some detail in Chapter 3, since this is an issue that will be investigated empirically in Chapters 5 and 6.

Sentence tense also appears to influence results in a somewhat inconsistent manner. The majority of studies investigating the ACE use present tense stimuli, after Glenberg and Kaschak (2002). A direct comparison between present and past tense sentences found the ACE in present tense stimuli only (Bergen & Wheeler, 2010). However, Aravena et al. (2010), using handshapes rather than response direction, found an ACE-type effect in Spanish past tense sentences. A recent fMRI study showed sensorimotor activation in response to Hebrew sentences in the past and present tense, but not to future tense sentences (Gilead, Liberman, & Maril, 2013). The within-study contrasts in which researchers alter the linguistic context (e.g., person; tense) while maintaining the same task, with diverging results, demonstrate that linguistic context does play a role in the interaction between language and action. However, the between-study inconsistencies (e.g., some authors finding

Table 2-1. Summary of effects from the embodied action semantics literature.

| <i>Author</i> | <i>Date</i> | <i>Journal</i> | <i>Stimuli</i> | <i>Task</i> | <i>Observations per cell</i> | <i>Direction of effect</i> |
|----------------------|-------------|--------------------------------|----------------|------------------------------------|----------------------------------|---|
| Ambrosini et al. | 2012 | <i>Conscious Cogn</i> | Word pair | Go/ no go sensibility judgement | 210 | Facilitation |
| Aravena et al. | 2010 | <i>PLOS ONE</i> | Sentence | Sensibility judgement | 676 | Facilitation |
| Bergen & Wheeler | 2010 | <i>Brain Lang</i> | Sentence | Sensibility judgement | 1400/1400 | Facilitation / null effect depending on gram. aspect |
| Bidet Ildei et al. | 2011 | <i>Acta Psychol</i> | Word | Task irrelevant word | 864 | Facilitation |
| Borghi | 2004 | <i>Acta Psychol</i> | Sentence | Part verification | 228 | Facilitation |
| Borghi et al. | 2004 | <i>Mem Cognition</i> | Sentence | Part verification | 760 | Facilitation |
| Borghi & Riggio | 2009 | <i>Brain Res</i> | Sentence | Sentence-picture matching | 1440 | Facilitation |
| Borghi & Scorolli | 2009 | <i>Hum Movement Sci</i> | Vb phrase | Sensibility judgement | 912/528 | Facilitation / null effect depending on blocking |
| Borreggine & Kaschak | 2006 | <i>Cognitive Sci</i> | Sentence | Go/ no go sensibility judgement | 240/240 | Facilitation / null effect depending on timing |
| Boulenger et al. | 2006 | <i>J Cognitive Neurosci</i> | Word | Go/ no go lexical decision | 378/378 | Interference / facilitation depending on timing |
| Boulenger et al. | 2008 | <i>J Physiology Paris</i> | Word | Task irrelevant word | 1750 | Facilitation |
| Bub et al. | 2008 | <i>Cognition</i> | Word | Lexical decision Silent reading | 768 | Facilitation / null effect depending on task |
| Buccino et al. | 2005 | <i>Cognitive Brain Res</i> | Sentence | Go/ no go semantic categorisation | 400 | Interference |
| Creem & Proffitt | 2001 | <i>J Exp Psychol Human</i> | Word | Word recall | 270 | Interference |
| De Vega et al. | 2013 | <i>Psychol Res</i> | Sentence | Silent reading | 420 | Interference |
| Dalla Volta et al. | 2009 | <i>Exp Brain Res</i> | Word | Go/ no go semantic categorisation | 160 | Interference |

| | | | | | | |
|------------------------|------|----------------------------|----------|---|---------|--|
| Diefenbach et al. | 2013 | <i>Front Psychol</i> | Sentence | Go/ no go sensibility judgement | 160/80 | Interference / facilitation depending on timing |
| Gentilucci & Gangitano | 1998 | <i>Eur J Neurosci</i> | Word | Task irrelevant word | 84 | Facilitation |
| Gentilucci et al. | 2000 | <i>Exp Brain Res</i> | Word | Task irrelevant word | 84 | Facilitation |
| Gianelli et al. | 2011 | <i>PLOS ONE</i> | Sentence | Sensibility judgement | 680/680 | Facilitation / null effect depending on subject |
| Glenberg et al. | 2002 | <i>Psychon B Rev</i> | Sentence | Sensibility judgement | 352 | Facilitation |
| Glenberg et al. | 2008 | <i>Q J Exp Psychol</i> | Sentence | Sensibility judgement | 352 | Facilitation |
| Glover et al. | 2002 | <i>Exp Brain Res</i> | Word | Task irrelevant word | 60 | Facilitation |
| Glover et al. | 2004 | <i>Exp Brain Res</i> | Word | Silent reading | 120 | Facilitation |
| Helbig et al. | 2006 | <i>Exp Brain Res</i> | Word | Picture naming | 288 | Null effect |
| Kaschak & Borreggine | 2008 | <i>Q J Exp Psychol</i> | Sentence | Action execution | 540 | Facilitation / null effect depending on timing |
| Kaup et al. | 2010 | <i>Brain Lang</i> | Sentence | Sensibility judgement | 160/160 | Facilitation / null effect depending on gram. aspect |
| Klatzky | 1989 | <i>J Mem Lang</i> | Sentence | Sensibility judgement | 128 | Facilitation |
| Liepelt et al. | 2012 | <i>Psychol Res</i> | Word | Action verb naming | 1280 | Facilitation |
| Lindemann et al. | 2006 | <i>J Exp Psychol Human</i> | Word | Go/ no go lexical decision/ letter detection / semantic categorisation/ | 576/300 | Facilitation / null effect depending on task |
| Masson et al. | 2008 | <i>Q J Exp Psychol</i> | Sentence | Reading aloud | 270 | Facilitation |
| Mirabella et al. | 2012 | <i>PLOS ONE</i> | Word | Go/ no go semantic categorisation | 180 | Interference |
| Myung et al. | 2006 | <i>Cognition</i> | Word | Lexical decision | 442 | Facilitation |

| | | | | | | |
|---------------------|------|-----------------------------|-----------|--|---------|--|
| Nazir et al. | 2008 | <i>Q J Exp Psychol</i> | Word | Go/ no go lexical decision | 378 | Interference |
| Postle et al. | 2013 | <i>Front Human Neurosci</i> | Sentence | Silent reading/ reading aloud | 675 | Interference |
| Rueschemeyer et al. | 2010 | <i>Neuropsychologia</i> | Word | Go/ no go lexical decision | 760 | Facilitation |
| Sato et al. | 2008 | <i>Brain Lang</i> | Word | Go/ no go lexical decision/ Go/ no go semantic categorisation | 120/120 | Interference/ null effect depending on task |
| Scorolli et al. | 2009 | <i>Exp Brain Res</i> | Sentence | Passive listening | 54 | Interference |
| Springer & Prinz | 2010 | <i>Q J Exp Psychol</i> | Word | Grammatical class judgement/ lexical decision | 918 | Facilitation |
| Taylor & Zwaan | 2008 | <i>Q J Exp Psychol</i> | Sentence | Silent reading | 816 | Facilitation |
| Taylor et al. | 2008 | <i>Brain Lang</i> | Sentence | Silent reading | 1248 | Facilitation |
| Tucker & Ellis | 2004 | <i>Acta Psychol</i> | Word | Task irrelevant word | 2560 | Facilitation |
| Van Dam et al. | 2010 | <i>Front Psychol</i> | Word | Go/ no go lexical decision | 1400 | Facilitation |
| Van Elk & Blanke | 2011 | <i>Acta Psychol</i> | Word | Iconicity judgement | 504 | Facilitation |
| Van Elk et al. | 2009 | <i>Cognition</i> | Word | Go/ no go semantic categorisation | 480 | Facilitation |
| Wang et al. | 2012 | <i>Acta Psychol</i> | Word | Spatial verification | 490 | Facilitation |
| Witt et al. | 2010 | <i>Psychol Sci</i> | Word | Picture naming | 1512 | Interference |
| Zwaan et al. | 2010 | <i>Brain Lang</i> | Narrative | Silent reading | 480 | Facilitation |
| Zwaan & Taylor | 2006 | <i>J Exp Psychol Gen</i> | Sentence | Silent reading/ sensibility judgement | 300 | Facilitation |

a null result for past tense sentences; other authors finding a significant ACE) mean that we cannot easily interpret the role of linguistic context.

2.4.2. Flexibility in simulation: Prediction versus experience

The increasing awareness of the importance of context in action-language cross-talk (see section 2.4.1. above) has led some researchers to view language-based motor resonance as a flexible process that, rather than simply replaying previous experiences, may in fact help us predict future events (Borghi, 2013; Rueschemeyer & Bekkering, 2013).

Until recently, most embodied accounts of language have stressed the role of past experience in motor resonance. For example, Barsalou's (1999) Perceptual Symbols account proposes that, during language comprehension, top-down simulators re-enact the state of activation that has previously been experienced when observing an object being referred to. Associationist theories argue that the relation between semantics and the motor system is based on Hebbian learning; during language acquisition, action verbs tend to be encountered in a motor context, leading to synaptic connections between the word and relevant motor area (Pulvermüller, Hauk, Nikulin, & Ilmoniemi, 2005; Pulvermüller, Shtyrov, & Hauk, 2009). When an action word is later encountered, the motor area is therefore immediately and automatically activated, without top-down intervention. In both cases, simulation is conceived as an internal non-conscious representation of *previous* actions. We call this process *retrospective simulation*. However, action simulation might also occur from a *prospective* point of view. In other words, simulation might involve forming predictive motor plans about future actions.

Prospective simulation and forward models

In the motor control literature, anticipation of upcoming actions is framed in terms of forward models, which compare predicted feedback from a given motor command with the actual feedback, and, in cases of conflict, signals to adjust the next motor command as

necessary (Miall & Wolpert, 1996; Wolpert, 1997). The motor system is therefore conceived as being engineered to anticipate the future actions, through a process of simulation (Wolpert & Flanagan, 2001), and indeed, one of the major benefits of simulation is that it allows anticipation of upcoming actions and their consequences (Prinz, 2006). Pickering and Garrod (2013) recently appealed to forward models as a means of integrating language production and comprehension, with a focus on communicative motor resonance. However, forward models might also influence language comprehension through referential motor resonance. In fact, Glenberg's Indexical Hypothesis (Glenberg & Kaschak, 2002; Kaschak & Glenberg, 2000) is based on combining affordances in action *plans*, rather than simple reenactments, and therefore hints at the possibility of prospective simulation.

Note that both retrospective and prospective simulation would draw on motor prototypes – that is, representations of how actions typically occur, based on previous experience and canonical affordances (Borghi & Riggio, 2009; Borghi, 2013). However, whereas retrospective simulation would simply involve running a simulation of a motor prototype, prospective simulation would involve using this prototype to predict future actions (see also Pezzulo & Castelfranchi, 2009). A key aim of this thesis is to investigate the relative roles of retrospective and prospective simulation on action language comprehension. Do language comprehenders merely simulate their previous long-term motor experience, or do they adapt to the current motor context in order to better predict future actions? One means of investigating this question is to look at body specificity in action language comprehension.

2.4.3. Body-Specificity

We saw in section 2.2.3 that motor cues seem to play a stronger role in action recognition than visual cues – people are better at recognising their own dancing movements than those of other people's, despite having more experience of observing other people dancing (Loula et al., 2005). But how specific are these self-based action representations? In

the next section we discuss evidence that action representations code for the specific effector (left or right hand) that an individual typically uses to perform particular actions.

The Body-Specificity Hypothesis

Left- and right-handed actors typically perform unimanual actions with their left or right hands, respectively. However, both groups have the experience of viewing other people perform more right-handed actions than left-handed actions, due to the larger number of right-handed actors in the world. The results of Loula et al.'s (2005) study suggest that the motor cue (dominant hand) should override the visual cue (right hand), and converging evidence suggests that this is indeed the case. For example, Sartori, Begliomini and Castiello (2013) recently used TMS to measure MEPs from the dominant and non-dominant hands of right- and left-handed participants as they observed a left- or right-handed grasping action. Participants showed increased MEPs in their dominant hand when observing both right- and left-handed actions, indicating that motor resonance reflects an individual's own motor experience in the world. Of course, we cannot rule out the possibility that participants initially activated a left-handed neural response to the left-handed stimuli, but then transferred this activation in order to prepare a right-handed response. Perhaps more compelling evidence comes from Willems, Toni, Hagoort and Casasanto (2009), who tested both left- and right-handed participants. In their study, participants imagined performing manual actions while undergoing fMRI; results showed differential activation between the two groups of participants, with right-handers showing left-lateralised activation, and left-handers showing right-lateralised activation. These results lend support to the Body-Specificity Hypothesis (Casasanto, 2011; Casasanto, 2009), which states that

“If concepts and word meanings are constituted in part by simulations of people's own perceptions and actions, then people with different bodily characteristics, who interact with their physical environments in systematically different ways, should form correspondingly different mental representations” (Casasanto, 2009, p. 351).

The Body-Specificity Hypothesis originated from a series of studies looking at emotional valence, in which participants were asked, for example, to indicate whether a positively or negatively valenced animal should be placed in a box on the right side of the page, or in a box the left side of the page (Casasanto, 2009). Previous research on right-handers had suggested that positive valence is associated with the right side of the body, and negative valence with the left side of the body (Davidson, 1992). However, when Casasanto tested both right- and left-handed participants, he found that right-handed participants displayed the expected right-is-good valence mapping, but that left-handed participants displayed the opposite left-is-good valence mapping. Further work has replicated this effect in children as young as 5 years old (Casasanto & Henetz, 2012). Brunyé and colleagues recently showed that the effects of body-specificity appear to extend to memory tasks as well as emotional valence: right-handed participants mis-remembered positively valenced locations as being further right, and negatively valenced locations as being further left, than was actually the case; left-handed participants showed the opposite pattern of results (Brunyé, Gardony, Mahoney, & Taylor, 2012).

Body-specificity and language

Given the posited link between action and action language semantics, and given evidence suggesting that representations of manual actions are body-specific, it is possible that semantic representations of action language might also be body-specific. In other words, right- and left-handed individuals may interpret action language differently from one another, based on their different motor experiences. To our knowledge, there have only been two studies looking at this question, and both studies used fMRI rather than behavioural paradigms. Willems, Hagoort and Casasanto (2010) tested for body-specificity in action language by having right- and left-handed participants perform a lexical decision task on manual action verbs (e.g. *throw*), with no response required on critical trials. When processing the manual action verbs, left-handed participants showed significantly more

activation in the right pre-motor hand area, and right-handed participants showed significantly more activation in the left pre-motor hand area. Therefore, it appears that a comprehender's experience of performing actions modulates their comprehension of isolated action verbs.

However, some caution is needed in interpreting these results, in light of a failure to replicate by Hauk and Pulvermüller (2011). Hauk and Pulvermüller were investigating language lateralisation for action verbs describing bimanual and unimanual actions. They recruited left- and right-handed participants to passively read the verbs. Although unimanual verb processing led to stronger activation in the left motor region for right-handers than for left-handers, activation in the corresponding right motor area was not significantly stronger for left-handed participants than for right-handed participants. Moreover, a later study by Willems and colleagues used TMS on the left and right premotor hand areas of right-handed participants, as they performed a lexical decision task on manual action verbs using left- and right-handed responses (Willems et al., 2011). Contrary to the authors' expectations, and the predictions of the Body-Specificity Hypothesis, there was no effect of which hand participants used to respond. It is clear therefore that more evidence is needed to establish to what extent right- and left-handed individuals interpret action language based on their different motor experiences: I will investigate this question in Chapter 5.

Plasticity in body-specificity

We have seen that the direction of effects in embodied approaches to language can vary depending on the context of the task (see section 2.4.1). But there is also some evidence that the preceding motor context can affect action-language comprehension. In one experiment, Borghi and Scorolli (2009) had right-handed participants perform a sensibility judgement on phrases describing actions typically performed either with the dominant hand (e.g., *to peel the apple*) or with the mouth (e.g. *to bite the apple*); participants responded using either their left or right hand. Participants were faster to respond to both mouth and

hand verbs using their right hand, reflecting the fact that the mouth actions typically also involved holding or lifting an object with the dominant hand. However, in a second experiment, participants performed sensibility judgements on phrases describing actions typically performed either with the dominant hand (e.g., *to throw the ball*) or with the foot (e.g., *to kick the ball*), participants were slower to respond to both types of verbs using their dominant right hand. Note that the foot trials included actions that could be performed with either (or both) foot (e.g., *to step on the grass*). The authors argue that the evocation of foot actions on half of the trials (at random), forced participants to adopt a general response strategy; in a later experiment, when hand and foot sentences were blocked, participants were faster to respond with their right hand on hand sentences only, as predicted by embodiment accounts. Therefore, it appears that the general motor context of a task can influence the interaction between motor resonance and action-language comprehension.

The fact that body-specificity appears at such a young age (Casasanto & Henetz, 2012) could be due to one of two possibilities: first, that the valence mapping is related to “built in” cognitive differences between left- and right-handers; or second, that the valence mapping may be acquired in a reasonably short time. In support of the second hypothesis, Casasanto and Chrysikou (2011) tested the left-right valence mapping of originally right-handed participants who, having suffered a stroke, had effectively switched their dominant hand in later life. Participants performed tasks from Casasanto’s original study (2009), and showed the same pattern of results (left-is-good) as the left-handed participants in that study. A further study showed just how quickly the valence mapping could be reversed: right-handed participants were randomly assigned to wear a bulky ski glove (in order to inhibit motor fluency) on their left or right hand while lining up dominoes (requiring a high level of manual dexterity). After only twenty minutes’ training, participants who wore the glove on their right hand showed the opposite pattern of valence mapping to those who wore the glove on their left hand. That is, those who had worn the glove on their right hand (and had thus been more fluent with their left hand in the training task), now mapped positively

valenced items to the left side of space. This flexibility in left- and right-handed participants' valence mapping suggests that their interpretations of action language may also display some plasticity, depending on the motor context of the task.

2.5. Summary and research questions

In this chapter, I provided an overview into the posited links between action and perception, and between action and language. I discussed evidence that *self* actions were more tightly connected to the sensorimotor system than *other* actions, and raised the possibility that first-person language might also be more tightly connected to the sensorimotor system than third-person language. In Chapter 3 of the thesis, I will explore issues of perspective taking, and first- versus third-person language in more detail. In Chapters 5 and 6, I investigate whether the posited differences between first- and third-person language are evident in comprehension. I also discussed body-specificity, and suggested that simulation might involve prediction and action planning, as well as motor experience. In Chapters 5 and 6 I test whether language comprehension interacts with long-term motor experience, short-term motor experience, or current motor plan. Finally, I discussed that fact that there are very few studies that investigate the effect of action on language, rather than the other way round. I explored some of the debate surrounding the possible functional role for motor resonance in language comprehension. In Chapter 7, I adapt a paradigm used by Klatzky et al. (1989) in order to test for a causal effect of action planning on language comprehension. Overall, the thesis attempts to specify some of the constraints that govern what happens when people understand sentences describing simple manual actions.

3. PERSPECTIVE TAKING IN LANGUAGE: INTEGRATING THE SPATIAL AND ACTION DOMAINS

3.1. Introduction

Over the past decade, research into language comprehension has increasingly been framed in terms of a link between perceptual and motor systems, and higher level cognitive tasks. A central assumption of such *Embodied Cognition* frameworks is that people's understanding of language is grounded in their physical interactions with the world (e.g., Barsalou, 1999, 2008; Fischer & Zwaan, 2008; Glenberg & Gallese, 2012; Glenberg, Sato, & Cattaneo, 2008; Pulvermüller, 2005). In strong versions of Embodied Cognition, language comprehension is achieved through mental representations that correspond, in perceptual or motor qualities, to the object or action being described. Such accounts draw on evidence that comprehenders are faster to correctly match sentences to images that correspond to the perceptual characteristics implied by the sentence context, such as orientation (Stanfield & Zwaan, 2001), shape (Pecher, van Dantzig, Zwaan, & Zeelenberg, 2009; Zwaan, Stanfield, & Yaxley, 2002), and implied movement (Kaschak et al., 2005, 2006).

In addition, Action-Sentence Compatibility effects (ACE; Glenberg & Kaschak, 2002) demonstrate that language comprehension is linked to action execution. Participants are faster to respond to sentences that imply moving the hand away from or towards one's body (e.g., "*Close/ Open the drawer*"), when the direction of response required (away from or towards their body) matches the direction of movement implied in the sentence. Aravena et al. (2010) recently provided evidence of a neural signature for ACE effects by recording event-related (brain) potentials. In this study, participants listened to sentences implying an open or closed hand shape, and indicated their understanding by responding with either an open or closed hand shape. Incongruent trials, where the hand-shape implied by the sentence

did not match the hand-shape required by the response, resulted in an N400 effect (associated with difficulty integrating stimuli into a given semantic context; Kutas & Federmeier, 2000). Such evidence is consistent with the viewpoint that action language comprehension involves representing an action as though you were performing it yourself – that is, from an agent’s perspective.

In this chapter, I explore research into action-perspective taking (from whose perspective do language users simulate a described action?), and spatial-perspective taking (from whose perspective do language users conceive spatial relations?). I propose that these two forms of perspective taking are fundamentally linked: in order for language users to perform an action simulation, they must first establish a spatial context for that action, by locating it within a situation model. In dialogue, spatial-perspective taking can be used by interlocutors to negotiate or align on situation models that specify similar spatial relations between entities, to ensure a mutually understood spatial context for actions. Actions are performed in space, and, therefore, we might expect considerable cross-over between the literatures on action- and spatial-perspective taking, but this does not appear to be the case. I argue that one reason for this situation is the use of inconsistent and conflicting terminology across the two fields.

One of the goals of this chapter is to unite action- and spatial-perspective taking in an account of action language comprehension. First, I propose a vocabulary for discussing action-perspective taking that will allow action- and spatial-perspective taking to be integrated. Next, I explore evidence from the Embodied Cognition literature, investigating which action-perspective comprehenders typically adopt. I argue that, contrary to some Embodied Cognition accounts where action-perspective taking is typically assumed to be fixed on the agent, several other perspectives are in fact available. I then review research into which spatial-perspective people tend to adopt in language use, and how such perspective taking is negotiated in dialogue. Finally, I propose the Spatial Grounding Hypothesis, which states that action simulations are grounded in spatial context. I discuss the

evidence in favour of this hypothesis, and explore the role of situation models in providing this context.

3.2. Representing other people's actions

At the same time as theories of action language processing have stressed the primacy of motor representations, theories of action understanding have argued that the same mental representations are involved in both performing and perceiving actions (e.g. Grèzes & Decety, 2001; Prinz & Hommel, 2002). For example, Common Coding theory (Hommel et al., 2001; Prinz, 1997) proposes that codes for planned actions and perceived actions share a common representational domain. In support of this account, behavioural research suggests first, that participants are less able to perceive a static stimulus (left or right pointing arrow) when performing a congruent action (left or right button press; Müsseler & Hommel, 1997), and second, that perceiving an action while planning an incompatible action affects action execution (Brass et al., 2000; Kilner et al., 2003). In other words, the link between perception and action affects our ability both to *perceive* stimuli, and to *perform* actions. Such findings are echoed by recent neurological research showing evidence of “mirror matching”, where regions of the motor system that are activated when performing an action are also activated when passively perceiving an action (e.g., Buccino et al., 2001; Grèzes, Armony, Rowe, & Passingham, 2003; for a review see Rizzolatti & Craighero, 2004).

Much research has argued that the perceiver of an action mentally simulates executing that action herself (Decety, 2002). This *simulation theory* has counterparts in simulation theories of mind that propose that understanding another person involves simulating their mental activity (e.g. Gallese & Goldman, 1998). Indeed, it could be argued that a successful theory of mind is one that allows us to predict and understand our own and other peoples' actions, and that this is achieved through simulation (Ruby & Decety, 2001). The close link between self and other then raises the question: how do we distinguish our

own actions or mental activities from those of other people? The ability to distinguish ourselves from other people is critical to successful social interaction, but in a system in which our own actions share representations with the actions of other people, action attribution becomes a key computational problem (de Vignemont & Haggard, 2008; Decety & Sommerville, 2003).

The mechanism by which the separation of self and other is maintained is beyond the scope of this paper (see, for example, Decety & Sommerville, 2003; Ruby & Decety, 2001; Ruby & Decety, 2004). But however it is achieved, the self-other distinction is tightly connected with perspective taking. First, self must be successfully distinguished from other in order for there to be the possibility of different perspectives (Jeannerod, 2006). Second, the ability to represent other people's actions in a similar way to their own allows people to take an agent's perspective on an action, even when they are describing or hearing about an action performed by somebody else.

3.3. *A taxonomy of perspective*

As highlighted above, a large body of research now suggests a link between language processing and sensorimotor activation (see Kiefer & Pulvermüller, 2012; Meteyard, Cuadrado, Bahrami, & Vigliocco, 2012 for recent reviews). This link can best be captured by Embodied Cognition accounts of language processing. Embodied Cognition seeks to distinguish itself from "traditional" psycholinguistic accounts by insisting that language representations are modal rather than amodal (e.g., Barsalou, 2008; Zwaan & Taylor, 2006). What is often not made explicit in Embodied Cognition accounts is that modal representations are inherently perspective-based. For a representation to be modal, it must assume a given perspective. In other words, the perspective is necessary to ground the representation. However, discussion of perspective taking in action language is often opaque, and this is particularly problematic if we wish to relate action-perspective taking and spatial-perspective taking.

In visual cognition, researchers distinguish between two types of spatial-perspective taking. Level 1 perspective involves understanding what falls within another individual's line of sight – for example, is a particular object occluded by another object as that person looks at it? Level 2 perspective involves understanding how the world appears from another person's perspective – for example, is a particular object to the left or the right of another object as that person looks at it? (Flavell, Everett, Croft, & Flavell, 1981; Michelon & Zacks, 2006). In the present paper, I limit the review of spatial-perspective to this second level, focusing on spatial relations, rather than visibility. Kessler and Rutherford (2010) argued that Level 2, but not Level 1 spatial-perspective taking, appears to involve some form of covert mental rotation or simulation. As such, Level 2 spatial-perspective entails a level of embodiment that Level 1 does not, and is therefore closer to the perspective-bound simulations proposed by Embodied Cognition accounts of action language understanding.

With respect to Level 2 spatial-perspective taking, we can contrast intrinsic, absolute, and relative reference frames (see Levinson, 1996, 2003). In an intrinsic reference frame, the position of an object is described relative to a reference object (e.g., “The window is above the door”). In an absolute reference frame, the position of an object is described in terms of stable environmental features, such as points of the compass, as in “The ship is south of the island”. Neither of these reference frames locates an object relative to an observer. A relative reference frame, on the other hand, does just that: for example, “The car is to my left”. Within a relative reference frame, one can adopt an *egocentric* or *allocentric* perspective. An *egocentric* perspective entails representing objects in a scene from your own viewpoint, and an *allocentric* perspective entails representing objects from the viewpoint of someone other than yourself (see Levinson, 2003 for a fuller treatment of spatial reference frames). The terms *egocentric* and *allocentric* therefore have specific and well-established meanings in the spatial literature: *egocentric* means conceptualising space from your own point of view, and *allocentric* means conceptualising space from another's point of view. In the literature on Embodied Cognition, however, researchers often use *egocentric* to refer to

putting oneself in someone else's shoes (for example, interpreting a sentence such as "John kicked Mary" as though the comprehender herself were performing the act of kicking; e.g., Willems, Hagoort, & Casasanto, 2010). This use of the term is opposite that in spatial-perspective taking and is therefore confusing. In addition, using the term *egocentric* perspective in action language, or *allocentric* perspective in spatial language, does not specify *whose* shoes the comprehender is putting herself into. In spatial language, this underspecification is typically not problematic, since the perspective adopted in a sentence such as "John is looking at the picture on the left" can be explicitly clarified. The comprehender can legitimately ask "on whose left?", and the speaker can reply "on *my* left", "on *your* left", "on *his* left", etc.

However, in action language, perspective-taking is implicit, rather than explicit, and no such clarification is possible. For example, a comprehender who responded to the sentence "John is looking at the picture on the left", with the query "who is looking?" would receive the reply "John", and remain no clearer about whose perspective the speaker was adopting. Therefore, unlike spatial language, when discussing action language it is necessary for embodied accounts to specify whose perspective is being adopted for a particular action: the term *egocentric* perspective tells us that comprehenders are putting themselves in somebody else's shoes, but crucially not whose shoes. Similarly, researchers often speak of "situated simulations" (Marino et al., 2012), or "sensorimotor experience" (Pecher et al., 2009) without specifying from whose perspective this simulation or resonance occurs. I suggest that this lack of specification derives from a widely held assumption in embodied cognition accounts that the agent's perspective is adopted. However, I also suggest that this assumption is unwarranted.

There are in fact different Embodied Cognition accounts of language processing, and researchers in this field place varying importance on the role of sensorimotor processing in semantics (see Meteyard et al., 2012 for a recent review of positions advocating different degrees of embodiment). However, a prevailing view conceives language comprehension as

an internal simulation of the described action, as if the comprehender were performing that action herself (e.g., Barsalou, 1999, Bergen & Wheeler, 2010; Borghi & Scorolli, 2009; Zwaan & Taylor, 2006). If it is true that action-perspective taking is fixed on the agent's perspective, then the underspecification of *egocentric*, outlined above, is not a problem; the perspective adopted would always coincide with the agent of the described action. However, as we shall see, it is not clear that an agent's perspective is always adopted. Researchers in action language therefore need to make clear exactly whose perspective they assume is being adopted.

For example, in understanding "John kicked Mary", there are at least two embodied perspectives that could be adopted for the action of kicking: that of John (the embodied agent); and that of Mary (the embodied patient). If the comprehender has reason to believe that other people are witnesses to the event (i.e., if she has reason to include bystanders in her situation model), then she can also adopt the perspective of a bystander watching the kicking event unfold (the embodied observer). For example, if a previous sentence implied the existence of a crowd gathering around Mary and John, the comprehender can adopt the perspective of a member of this crowd, observing John kicking Mary. In each case, the comprehender represents the action from the perspective of a person present in the comprehender's model of that event. In taking the embodied agent's perspective, the comprehender represents the action of kicking as though she herself were the agent of that action, by activating the same systems involved in executing a kicking action. In taking the embodied patient's perspective, the comprehender represents the action of kicking as though she herself were the patient of that action (presumably activating some form of empathic response to the pain, such as wincing). In taking the embodied observer's perspective, the comprehender represents that action as though she were watching it unfold, by activating the same systems that would be recruited when observing such an action.

In addition to these embodied perspectives, there is another perspective that the comprehender could take: that of the non-embodied observer. Unlike an embodied

participant or observer, the non-embodied observer represents the action from the perspective of someone who is not included in the comprehender's model of that event. This can be made clear by contrasting the sentence "John kicked Mary" with the sentence "John kicked Mary when they were alone": only in the first sentence is it possible to adopt an embodied observer's perspective. I suggest that comprehenders may adopt the perspective of a non-embodied observer when the spatial context is insufficient to allow an embodied action simulation.

The sentence "John kicked Mary" refers to a transitive event with two participants. There are of course, more complex sentences in which further embodied perspectives exist. This is the case for sentences describing ditransitive events (e.g., "John passed the child to his wife"), or sentences where a thematic role is occupied by more than one entity (e.g., "John kicked Mary and Sam"). The number of potential embodied perspectives available for a given sentence is therefore the number of participants in that event plus that of any embodied observers licensed by the comprehender's situation model. I propose that these perspectives (e.g., embodied agent, embodied patient, embodied recipient..., plus embodied observer and non-embodied observer) provide a transparent basis for discussing action perspective taking. Using these terms, researchers can not only distinguish between embodied and non-embodied representations, but within the embodied representations, it is possible to distinguish whose perspective is adopted.

3.4. Do language users consistently adopt the agent's perspective?

I noted above that many embodied accounts of language assume that if a perspective is adopted for action language, it is the agent's perspective (e.g., Glenberg & Kaschak, 2002; Wu & Barsalou, 2009; Zwaan & Taylor, 2006). Such an assumption is consistent with results from studies using isolated action verbs, for example, showing somatotopic activation

for specific body parts. Research using functional magnetic resonance imaging (fMRI) has found that passive listening to an arm-word (“pick”) leads to increased activation in areas of the premotor and primary motor cortex associated with arm movements; passive listening to a face-word (“lick”) leads to increased activation in areas associated with the face; and passive listening to a foot-word (“kick”) lead to increased activation in areas associated with the feet (Hauk, Johnsrude, & Pulvermüller, 2004; see also Aziz-Zadeh, Wilson, Rizzolatti, & Iacoboni, 2006). In other words, the activation appears to be associated with particular acts from the perspective of the agent of the act (e.g., the kicker) rather than (for example) the patient (e.g., the person or thing that is kicked). Further work using magnetoencephalography (MEG) has demonstrated that such somatotopic activation occurs extremely quickly, within 200 ms of word presentation, and even when participants are concentrating on an unrelated, non-language based task (Pulvermüller, Shtyrov, et al., 2005). These findings suggest that adopting an embodied agent’s perspective may occur automatically in the early stages of semantic processing, at least in isolated words.

More evidence that people adopt the embodied agent’s perspective (as though the comprehender herself were carrying out an action) comes from evidence for “body-specific” representations of manual action verbs (e.g. *throw*) in a Dutch lexical decision task (Willems, Hagoort, & Casasanto, 2010). Left-handed participants showed activation in the right pre-motor hand area, but right-handed participants showed activation in the left pre-motor hand area, despite there being no manual responses on critical trials. These results echo findings of “body-specific” activation for motor imagery, where left- and right-handed participants imagined performing actions described by manual action verbs (Willems et al., 2009). It therefore appears that people tend to adopt the embodied agent’s perspective for isolated verbs, representing the verb according to how they personally would perform those actions with their particular bodies (i.e., right-handed for right-handed participants; left-handed for left-handed participants).

However, verbs are usually processed not in isolation, but in the context of sentences featuring noun phrases that refer to particular entities. Do language users also adopt an embodied agent's perspective in action sentences, as well as isolated verbs? The evidence that they do is mixed. Participants undergoing fMRI were presented with Italian versions of mouth-, leg-, or hand-related action sentences featuring the first-person in the agent's role (e.g., "I bite the apple"; "I grasp the knife"; "I kick the ball"; Tettamanti et al., 2005). The results showed evidence of somatotopic activation similar to that observed in isolated verb processing (e.g., Hauk et al., 2004), implying that participants were simulating the described actions from the agent's perspective. However, in this study, the agent's perspective coincided with the perspective of the potentially self-referential pronoun "I": participants may have adopted a perspective in line with the thematic role assigned to the pronoun "I", rather than the perspective of the agent per se. A better indication of whether participants routinely adopt the embodied agent's perspective comes from studies investigating ACE effects (Glenberg & Kaschak, 2002; Glenberg, Sato, Cattaneo, et al., 2008). When sentences were given in the form of an imperative (e.g., "Close the drawer"), participants were faster to respond when the direction of the response was congruent with the movement implied by the agent in the sentence than when it was incongruent. In other words, they appeared to adopt the perspective of an agent closing a drawer. However, in sentences featuring two arguments, one of whom could refer to the participant, participants were faster to respond when the direction of the response was congruent with the movement relative to the pronoun "you". For example, participants were faster to respond with away movements to sentences such as "You delivered the pizza to Andy", but faster to respond with towards movements to sentences such as "Andy delivered the pizza to you". Therefore, this suggests that when a sentence involves a potentially self-referential pronoun ("you", "I"), comprehenders tend to adopt the perspective of the thematic role assigned to that pronoun, whether or not this coincides with the thematic agent of the action. In a dialogue context, where sentences such as "You are / I am slicing the tomato" are uttered and

understood by each participant in turn, the situation is more complex. Participants appear to prioritise adopting opposing perspectives for “you” and “I”, over maintaining a consistent perspective (e.g. embodied agent, embodied observer) for either of the pronouns (Pickering, McLean, & Gambi, 2012).

Several studies have addressed whether people adopt the agent’s perspective when the agent of a described action is not self-referential, in the absence of a second self-referential argument. In Embodied Cognition accounts that conceive action language as an extension of mirror-matching, where representations of other people’s actions are inherently similar to representations of one’s own actions (e.g., Pulvermüller, 2005; Rizzolatti & Arbib, 1998), descriptions of actions performed by third-person agents should elicit similar effects to descriptions of actions performed by first- or second-person agents. In line with this prediction, Buccino et al. (2005) used transcranial magnetic stimulation (TMS) to stimulate the left-hemispheric hand or foot motor areas, as participants listened to Italian third person hand- or foot-related action sentences (e.g. “He sewed the shirt”; “He marched on the spot”), compared with control abstract sentences (e.g. “He loved his wife”). Motor evoked potentials (MEPs) from the hand and foot muscles were recorded. Hand MEPs were modulated specifically when listening to hand-related action sentences, and foot MEPs were modulated specifically when listening to foot-related sentences. These results suggest at least some tendency to adopt an embodied agent’s perspective for third-person sentences.

However, without a direct comparison between first- and third-person sentences, we cannot know whether action perspective-taking in third-person sentences matches action perspective-taking in first-person sentences. Behavioural evidence suggests that comprehenders reading self-referential and non-self-referential sentences adopt different action-perspectives. Brunyé, Ditman, Mahoney, Augustyn, and Taylor (2009) used a sentence-picture matching task with first-, second-, and third-person action sentences, and “internal” or “external” action images. In the “internal” images, the position of the hands meant they could plausibly be interpreted as those of the participant. In the “external”

images, the position of the hands meant they could not plausibly be interpreted as those of the participant. Instead, they could most plausibly be interpreted as those of an agent who the participant was observing perform the action. Selecting an internal image would imply adopting the embodied agent's perspective. Selecting an external image would imply adopting the perspective of an embodied observer. Brunyé et al. found that participants were faster to correctly match first- and second-person sentences to internal rather than external images, and to correctly match third-person sentences to external rather than internal images. In other words, participants adopted the embodied agent's perspective when the agent of the sentence could be attributed to the comprehender, but not otherwise (see also Ditman, Brunyé, Mahoney, & Taylor, 2010; Sato & Bergen, 2013).

In an fMRI study, Tomasino, Werner, Weiss and Fink (2007) found no difference in primary motor cortex activation between silent reading of German action phrases presented in the first-person (e.g., "I hammer") versus third-person (e.g., "he hammers"). However, Papeo, Corradi-Dell'Acqua, and Rumiati (2011) had participants silently read action or non-action Italian verbs conjugated in the first- or third-person (e.g., "I write"; "he writes"; "I wonder"; "he wonders"). They found that TMS-induced MEPs in the relevant motor area (e.g., hand) increased for the first-person action verbs, but that the third-person action verbs behaved like the non-action verbs, and showed no increase in MEPs. Embodied Cognition accounts need not predict total parity between first- and third-person action representations. However, the posited involvement of the motor system in action language comprehension (e.g., Fischer & Zwaan, 2008) should imply at least some difference between third-person action and non-action verbs. The fact that a difference between action and non-action verbs was found only in first-person sentences, led Papeo et al. (2011) to conclude that motor simulation of an action sentence occurs only when the self is identified as the agent of the action.

What could be behind the conflicting results of Tomasino et al. (2007), and Papeo et al. (2011)? One important difference may be in the task. Participants in Tomasino et al.'s

study were asked to decide whether a described event took place inside or outside a building, and thus could complete the task without paying attention to whether the verb was presented in the first- or third-person. On the other hand, Papeo et al. instructed participants to determine the syntactic subject of a phrase, thus focussing attention on the contrast between first- and third-person agents. Researchers are becoming increasingly aware of the role of task demands and context in studies of Embodied Cognition. The conflicting results here add to evidence suggesting that motor representations of action language may not be activated automatically, but depend on aspects of the task, including depth of processing (Sato, Mengarelli, Riggio, Gallese, & Buccino, 2008), sentence tense (Bergen & Wheeler, 2010), and relevance to task goals (Hoedemaker & Gordon, 2013). Indeed, it is possible to view the emphasis, outlined above, on the agent's perspective as a result of task demands. The link between action and language has typically been investigated by studying congruency effects when participants execute actions during sentence processing (Taylor & Zwaan, 2008; Zwaan & Taylor, 2006), after sentence processing (Glenberg & Kaschak, 2002; Glenberg, Sato, Cattaneo, et al., 2008), or before sentence processing (Glenberg, Sato, & Cattaneo, 2008). When the emphasis of the task is to execute an action, it is perhaps not surprising that results seem to indicate that participants adopt the agent perspective. Other paradigms in embodied approaches to language follow sentence processing with image presentation rather than action execution. For example, participants are typically faster and more accurate to recognise an image of an object when it is presented in the same orientation (vertical/horizontal) as implied by the preceding sentence (Stanfield & Zwaan, 2001; see also Pecher et al., 2009; Zwaan et al., 2002). The authors interpret these findings as evidence that comprehenders run visual simulations of an event (i.e., they adopt an embodied observer's perspective). The perspective adopted by comprehender may therefore depend on the task used to investigate it. It may even be possible to use the task to prime participants to adopt a given action-perspective, although we know of no study that has investigated this possibility.

In summary, some Embodied Cognition accounts of action language assume that people adopt an embodied agent's perspective when comprehending action language, based on an internal simulation of performing that action (Barsalou, 2009; Zwaan & Taylor, 2006). Moreover, strong Embodied Cognition accounts assume that the agent's perspective is automatically activated, regardless of contextual factors such as the reference of the sentence, as determined, for example, by the subject pronoun (Pulvermüller, Shtyrov, et al., 2005; Pulvermüller, 2005). The evidence outlined above suggests that people do adopt the embodied agent's perspective for isolated verbs, and for sentences in which a potentially self-referential pronoun ("you", "I") is specified as the agent (Hauk et al., 2004; Pulvermüller, Shtyrov, et al., 2005; Willems, Hagoort, et al., 2010). However, when a self-referential pronoun occupies a thematic role other than agent, comprehenders appear to adopt the perspective of the thematic role assigned to that pronoun, and not the perspective of the agent (Glenberg & Kaschak, 2002). When a third party is specified as the agent of an action, and no self-referential pronoun is present, some evidence suggests that comprehenders adopt the embodied agent's perspective (Buccino et al., 2005; Tomasino et al., 2007), whereas other evidence suggests that people adopt an embodied observer's perspective (Brunyé et al., 2009; Papeo et al., 2011). Although more data are clearly needed in order to draw firm conclusions about which perspective comprehenders adopt under which circumstances, current data demonstrate that adopting an agent's perspective is not the only possibility during action language comprehension. As a consequence, the underspecified terms *egocentric* or *internal* perspective should be avoided when discussing action-perspective taking. Instead, researchers in Embodied Cognition should seek to employ more transparent terms that specify *in whose shoes* the comprehender is placing herself (e.g., *embodied agent*, *embodied patient*, *embodied observer*).

3.5. *Spatial perspective-taking*

So far, I have reviewed evidence examining whose action-perspective language users tend to adopt when processing action language sentences. However, language users can also adopt a range of spatial-perspectives during language production or comprehension. Of particular interest is whether people adopt an egocentric spatial-perspective (conceiving spatial relations from their own point of view), or an allocentric spatial-perspective (conceiving spatial relations from another's point of view).

Schober (1993) asked participants to describe the location of objects, either alone, to an imaginary addressee, or when in the same room as a conversational partner. Participants were more likely to describe the location from the addressee's point of view, using terms such "on your left", than from their own point of view. Schober (1995) also found that speakers tended to adopt the addressee's perspective in task requiring the speaker to identify particular objects to an addressee. Interestingly, participants in Schober (1993) who described objects to an imaginary addressee were *more* likely to use the addressee's perspective than participants whose conversation partners were present. With an addressee absent and unable to provide feedback, it may be safer for the speaker to assume the addressee's perspective as often as possible. Duran, Dale, and Kreuz (2011), using a virtual reality paradigm, also found that participants were more likely to adopt an allocentric spatial perspective when told that they were interacting with a virtual, rather than real partner. It appears that believing that their partner was real allowed participants to shift more of the burden of mutual comprehension to their partner. The tendency to shift responsibility for effective communication to a conversation partner may be stronger when, as in Duran et al.'s study, that partner is making a request rather than providing information. Yoon, Koh, and Brown-Schmidt (2012) found that speakers in a modified referential communication task were more likely to use allocentric perspective when requesting something from their partner compared with giving information to their partner. Since it is in speakers' interests to

ensure that their requests are successfully understood, it is sensible for listeners to assume that speakers will adopt an allocentric perspective when making that request.

The above results show that spatial-perspective taking, like action-perspective taking, is a flexible process. By changing the perspective they adopt, speakers or listeners can shift more or less of the burden of mutual comprehension on to their partner. Further research suggests that during dialogue, people may attempt to minimize not only their own effort, but the collective effort of both conversation partners, by obeying what Clark and Wilkes-Gibbs (1986) term the principle of *least collaborative effort*. Speakers and listeners often appear to adopt spatial perspectives in a way that maximises the resources available. The principle of least collaborative effort appears to be adopted especially in cases where one partner is judged less able to complete the communication task (Schober & Brennan, 2003). For example, Mainwaring, Tversky, Ohgishi, and Schiano (2009) found that speakers were more likely to use an (allocentric) addressee's perspective when the addressee was under increased cognitive load. Schober (2009) studied what happens when, unbeknownst to the participants, one partner in a conversation has better spatial ability than another, as determined by mental rotation test results. Participants were paired into a director and a matcher, with no knowledge of their own or their partner's results on the mental rotation tests. The matcher selected a target circle from an array, based on the director's spatial descriptions. Low-ability directors were more likely to take their own (egocentric) perspective, while high-ability directors were more likely to take their partner's (allocentric) perspective. Over the course of the experiment, high-ability directors who were paired with low-ability matchers increased their use of allocentric perspective, whereas low-ability directors who were paired with high-ability matchers decreased their use of allocentric perspective. Note that these opposite patterns of behaviour between high- and low-ability directors is in itself reason to be cautious of basing our understanding of spatial perspective-taking in language on university students of (presumably) high cognitive ability.

I suggest that this online adaptation to a partner's ability to engage in the communicative task is compatible with conversation as conceived as a joint action (Clark, 1996; Gambi & Pickering, 2011; Sebanz, Bekkering, & Knoblich, 2006). In the case of spatial perspective-taking, the perspective that people adopt appears to depend at least partly on the ability of their partner to engage in the task. In the next section, I argue that maximising the collective resources in this way allows conversation partners to establish coherent situation models in both partners. Once these situation models have been established, language users are in a position to adopt a particular action-perspective when performing mental simulations of actions. However, interlocutors do not adapt only their use of spatial-perspective within a relative reference frame; they also appear to adapt their choice of reference frame itself. Evidence that conversation partners align on their use of reference frame comes from studies using a confederate-priming paradigm. Watson, Pickering, and Branigan (2004) studied participants' use of an intrinsic versus a relative reference frame. Participants were more likely to use an intrinsic reference frame after the confederate had used an intrinsic frame than after the confederate had used a relative reference frame. Importantly, Watson et al. found participants regularly switched between reference frames. Spatial-perspective taking in dialogue is therefore highly flexible in order to allow for maximal alignment and hence maximal similarity in situation models. Whether such alignment on situation models occurs as a result of automatic priming (e.g., Pickering & Garrod, 2004, 2006), or of negotiating common ground (e.g., Clark, 1996) is beyond the scope of this chapter, but I assume both possibilities remain open.

3.6. Situation models: Linking spatial- and action-perspectives

Much research on Embodied Cognition can be traced back to studies of situation models in language processing (e.g. Johnson-Laird, 1983; Van Dijk & Kintsch, 1983).

According to recent accounts, situation models are representations of specific situations described in language, where events are connected along five dimensions; space, time, protagonist, causality, and intentionality (Zwaan, Langston, & Graesser, 1995; for a review of situation models in language see Zwaan & Radvansky, 1998). Evidence suggests it is the content of these models, rather than linguistic form of the language itself, which is typically retained in memory and integrated into updated models as comprehension continues (Johnson-Laird & Stevenson, 1970; Sachs, 1967). For example, Bransford, Barclay, and Franks (1972) demonstrated that participants who read the sentence “Three turtles rested on a floating log, and a fish swam beneath them” frequently selected the linguistically different but situationally equivalent sentence “Three turtles rested on a floating log, and a fish swam beneath it” in a recognition test (see also Barclay, 1973; Honeck, 1973; but see Jahn, 2004 for potential limits on such effects). Many modern studies in the Embodied Cognition literature have found similar effects when the focus is shifted to online rather than memory processes. For example, Borghi, Glenberg, and Kaschak (2004) found that participants were faster to verify items typically found inside a given object (e.g., “steering wheel”) following a preamble placing them inside that same object (e.g., “You are driving a car”) versus outside it (e.g., “You are refuelling a car”). They proposed that participants used a mental simulation grounded in modal representations (e.g., of being inside or outside a car), which then guides property verification (see also Kosslyn, Ball, & Reiser, 1978).

Such mental simulations are a defining feature of embodied theories of language, and differ from the situation models discussed in text or discourse processing in that they appear to capture online processing during language comprehension. Whereas situation models represent the integration of knowledge about events and situations into a coherent, existing framework, mental simulations are concerned with the online action-perspective taking about a particular act (see also Zwaan, 2008 for discussion of the differences). I propose that this “nesting” of action simulations within situation models is what links spatial- and action-perspective taking in language. In order for a comprehender to adopt an

embodied perspective on an action, that action must be grounded in a spatial context. This spatial context is provided by the comprehender's situation model. Situation models are conceived from a particular spatial perspective; in dialogue, conversation partners maximise their resources and align on spatial-perspective and reference frames, in order to ensure suitably similar situation models, for example by making use of the principle of least collaborative effort (Clark, 1996). Recall that situation models can specify events across a number of dimensions (space, time, causality, etc.; Zwaan et al., 1995). For our purposes, "suitably similar" situation models means that the situation models of both interlocutors specify the same protagonists in roughly the same spatial relations to one another.

The spatial relations between objects and people are a fundamental part of situation models (Tversky, 1991), and might be specified at various levels of granularity, from coarse grained, specifying only overall direction, to fine grained, specifying exact distances. I propose that the minimum information required in a situation model in order to run an action simulation is the participants in that action and some (coarse-grained) information about the spatial relations in which they stand. This allows comprehenders to establish the direction and perhaps rough distance in which an action occurs, and thus to simulate it, adopting a particular action-perspective. When a sentence is interpreted self-referentially (because it involves pronouns such as "you" or "I" – and perhaps also, although we know of no study demonstrating this – when it refers to the comprehender by name), the comprehender creates a situation model grounded in his or her own body; other participants in the action are by default conceived as located in front of the comprehender. For example, in Glenberg and Kaschak (2002), sentences such as "You delivered the pizza to Andy" elicited ACE effects because the direction of an action could be established (away from the comprehender's body), and an action-perspective could be adopted in line with the thematic role assigned to the self-referential pronoun (embodied agent). I refer to the idea that spatial context grounds action-perspective taking as the Spatial Grounding Hypothesis.

The Spatial Grounding Hypothesis can explain the diverging results we discussed earlier regarding first-person and third-person language. Recall that Papeo et al. (2011) found that comprehenders appeared to adopt an embodied agent's perspective for first-person language, but no embodied perspective for third-person language; whereas the results of Tomasino et al. (2007) suggested that first- and third-person language elicited similar action perspectives. The Spatial Grounding Hypothesis explains these results as follows. In Papeo's study, the first-person sentences ground the situation model in the comprehender's own body, allowing an action simulation to occur; in the third-person sentences, the situation model contains insufficient spatial information for action simulation. In Tomasino et al.'s (2007) study, the task was to decide whether the described action took place inside or outside, thus encouraging the construction of situation models in which to situate first- *and* third-person actions. Task demands may therefore play an important role in action language understanding, in the extent to which they provide, or encourage participants to create, spatial context for the described actions.

For example, third-person sentences in which the direction of the described action (e.g., turning a knob clockwise or anti-clockwise) is apparent from the sentence context (e.g., raising or lowering the volume) also elicit ACE-type effects where the comprehender adopts an embodied agent's perspective (Zwaan & Taylor, 2006). Further work suggests that these effects only occur once the direction of movement (clockwise or anti-clockwise) has been specified (Taylor et al., 2008). On the other hand, some evidence suggests that where a described action lacks suitable spatial grounding – for example, when it is described in the third-person, and the spatial relations between participants are not specified – action-perspective taking does not occur. Gianelli, Farnè, Salemme, Jeannerod, and Roy (2011) replicated the ACE effects in sentences featuring second-person agents (e.g. “You gave a pizza to Louis”), but not third-person agents (e.g. “Lea gave a pizza to Louis”). When avatars provided spatial locations for the third-person agents, the ACE effect reappeared. In

other words, participants only adopted an embodied agent's action-perspective when their situation model afforded adequate spatial context.

I have suggested that spatial context grounds action-perspective taking, such that a comprehender can only simulate an action from a particular perspective if her situation model specifies the participants in that action, and their spatial relations (thus giving her access to the direction in which an action would occur). I have argued that this proposal, the Spatial Grounding Hypothesis, can incorporate apparently conflicting results about action-perspective taking into a coherent framework. But there are other factors that support the Spatial-Grounding Hypothesis. First, it predicts that conversation partners will align on spatial-perspective and choice of reference frame, in order to establish similar situation models in both partners. We saw in the previous section that this is indeed the case. Second, it can explain why the presence of a potential agent other than the speaker affects how likely the speaker is to shift her spatial perspective. Tversky and Hard (2009) investigated the influence of a potential agent on how likely people were to adopt an allocentric perspective. Participants viewed photographs of scenes in which an actor was reaching for objects (and thus, in a position to act on that object), scenes with no actor, and scenes with an actor who was not reaching. Participants were more likely to adopt an allocentric spatial perspective (that of the actor in the photograph) when the actor was reaching versus not reaching for an object. Similarly, Zwickel (2009) investigated what spatial-perspective participants adopted when watching clips of animated triangles they perceived as more or less agentive (Abell, Happé, & Frith, 2000). Zwickel provided some evidence that participants only adopt an allocentric perspective when they view the other entity as an agent with specific states of mind, rather than a non-agentive entity moving at random. Mazzearella, Hamilton, Trojano, Mastromauro, and Conson (2012) recently extended Tversky and Hard's study by manipulating the extent to which the actor was in a position to act on the object (grasping versus gazing). Images in which the actor was in a better position to act on the object (grasping) triggered more use of allocentric spatial perspective in participants compared with

images in which the actor was in a less good position to act on the object (gazing). All of this suggests that participants are more likely to adopt an allocentric spatial-perspective in the presence of someone they perceive as a potential agent.

On the other hand, research suggests that the ability to extract information useful for object interaction (e.g., size) is diminished when participants adopt an allocentric, rather than egocentric spatial perspective (Campanella, Sandini, & Morrone, 2011). In addition, participants are faster to execute a reach-to-grasp movement when the object also falls within the peripersonal, rather than extrapersonal, space of a second person, implying that people tend to be faster to interact with objects in the presence of another potential agent (Gianelli, Scorolli, & Borghi, 2013). Given that participants want to interact with objects more quickly in the presence of another potential agent, and given that adopting an allocentric perspective may impede their ability to do so, why, then, would participants be more likely to adopt an allocentric perspective in the presence of another potential agent? Tversky and Hard suggested that their participants, in order to make sense of the scene, tried to understand the possibility that the other person can interact with the objects. We propose that people find it easier to understand another person's potential actions when they understand the spatial relations in the other person's situation model; that is, when they conceive space from that person's perspective. Spatial-perspective taking can therefore augment a situation model by increasing awareness of an agent's *potential* actions, even when no action is described.

One argument against the Spatial Grounding Hypothesis is that that situation models are often underspecified, and do not provide comprehenders with the necessary spatial context in which to situation action simulations. In particular, isolated verbs provide no explicit spatial context, and yet evidence suggests that comprehenders do adopt an embodied agent's perspective on the actions that the verbs describe (e.g., Hauk et al., 2004; Willems et al., 2010). I suggest that participants typically interpret these isolated verbs as self-referential (even when they are not presented in the imperative). Thus, like explicitly self-referential

language, the comprehender's own body grounds her situation model in this case. In other cases, where the comprehender's situation model does not allow her to establish at least the coarsely-coded spatial relations involved in an action, she cannot adopt an embodied action-perspective, because the action simulation cannot be run. However, this does not mean that the sentence describing an action cannot be understood. Rather, the comprehender can adopt the perspective of a non-embodied observer. This perspective is not an embodied perspective, in the sense that it does not involve a simulation of the action from the perspective of any of the participants. However, it is sufficient to allow the comprehender to understand the sentence, even if that understanding is somewhat less fully specified than the situation in which an embodied action perspective can be adopted. Researchers have found that non-ice hockey players respond more slowly and show less pre-motor activation than expert ice hockey players do when reading sentences about ice hockey (Beilock, Lyons, Mattarella-Micke, Nusbaum, & Small, 2008), but this does not mean that fail to understand the sentences. Their understanding may be impoverished relative to that of the expert players, but comprehension is not an all or nothing process (Taylor & Zwaan, 2013). Just as non-expert players may supplement their understanding of ice hockey using information and inferences about similar experiences (e.g., playing field hockey), comprehenders with inadequate situation models may supplement their models by adopting a non-embodied observer's perspective based on memories or inferences about similar situations.

Allowing for the possibility of a non-embodied perspective is important, since it means that embodied accounts are not bound to the claim that motor resonance is *necessary* for language comprehension. The necessity claim is problematic for embodied accounts of language processing for two reasons. First, because it implies that all language must be understood through motor activation, even when there is no obvious source of activation, such as in abstract and metaphorical language. Some proponents of embodiment argue that non-literal language can be understood through the transfer of motor activation in concrete cases to more abstract cases, perhaps through a process of neural exploitation (e.g., Gallese

& Lakoff, 2005). However, the evidence for motor activation during comprehension of metaphorical language is mixed at best (see section 8.3.4). Second, and more importantly, the claim that simulations are necessary for comprehension implies that participants who do not run a simulation (e.g., in cases where no ACE is observed – see for example Bergen & Wheeler, 2010; Gianelli et al., 2011) do not understand the sentences being presented, even though they are able to answer questions about those sentences to the same degree of accuracy as participants who do show evidence of simulation. The necessity claim therefore requires that comprehension involves something quite different from our intuitive grasp of what it is to understand a sentence (see Hickok, 2010, and section 2.2.1 for a similar argument in action perception). Thus, an account of sentence comprehension which argues for a necessary link between language and action leaves itself open to the accusation that it is in fact a theory of something other than sentence comprehension. I therefore argue that in order to remain plausible, embodied accounts of language must still be able to accommodate a non-embodied perspective. The Spatial Grounding Hypothesis provided a starting point for exploring under what circumstances comprehenders adopt this non-embodied perspective rather than running an action simulation (see also section 8.3.6).

3.7. Summary

In this chapter, I have attempted to reconcile two largely distinct literatures concerned with spatial-perspective taking and action-perspective taking. I have proposed a transparent vocabulary for action-perspective taking, which I hope will facilitate research between these two domains. At the heart of the proposal is the suggestion that researchers working in Embodied Cognition must specify *from whose perspective* a given action is being simulated. Although an agent's perspective seems in many cases the most natural candidate, other perspectives are possible, and are often adopted when self-referential pronouns are assigned a thematic role other than agent.

I have argued that comprehenders can only adopt an action-perspective if they have a spatial context for that action (the Spatial Grounding Hypothesis). In the case of isolated verbs and self-referential pronouns, people typically take their spatial grounding from their own bodies. But in the absence of self-referential language, action-perspective taking can only occur when the spatial relations between participants in the action have been established within the comprehender's situation model. In dialogue, interlocutors use spatial-perspective taking to ensure that each partner's situation model specifies similar spatial relations.

4. STATISTICAL ANALYSES USED IN THIS THESIS

The experiments in this thesis analysed the relationship between a variety of independent variables (e.g., dominant hand, response hand, planned action), and a set of dependent variables – both dichotomous (e.g., accuracy, image choice) – and continuous (e.g., button release RT, button press RT, voice onset time), using mixed-effects modelling. In modelling terms, these independent variables are fixed effects. Mixed-effects modelling allows us to combine fixed effects with random effects terms sampled from a larger population, such as participant or item (Baayen, 2008).

4.1. Mixed-effects modelling

The inclusion of participants and items as random effects allows linear mixed-effects (henceforth, LME) models to estimate the parameters by which different participants or items vary: some participants may respond slower overall than other participants, and some items may be more easily processed overall than others. Inclusion of random intercepts for participants and items allows us the model to capture this variance. Moreover including by-participant or by-item random slopes allows us to additionally account for the fact that some participants or items might perform differently in different experimental manipulations (Baayen, 2008). In addition, logit link LME models have been demonstrated to handle categorical data (e.g., image choice) better than ANOVA (Jaeger, 2008).

4.2. Maximal random effect structure

Recent years have seen increasing use of LME modelling in psycholinguistics research, and as such, recognised best-practice in model construction and reporting has changed as this thesis progressed. In particular, Barr, Levy, Scheepers and Tily (2013) argued that for hypothesis testing, maximal random effect structure (i.e., random intercepts for subjects and items, and random slopes for all within subjects and within items terms)

should be used wherever possible. One potential problem with this approach is that models with complicated random effects structure may not converge. I explain how this was dealt with in section 4.3 below. It should be noted that earlier analysis using forward model selection revealed the same pattern as results as reported here.

4.3. Model construction

In all analyses, I built a full model with all fixed effects, and maximal random effect structure (i.e., random intercepts, including interactions, for between-subjects factors; and random intercepts and random slopes, including interactions, for within-subjects factors; Barr, Levy, Scheepers, & Tily, 2013). Where the resulting model failed to converge, I simplified the random effects structure as specified in the text through backwards selection. Where simplification of random effect structure was necessary, I report the best-fitting of the converged models. Model fit was assessed by comparing model fit with and without the relevant random effect using the log-likelihood ratio χ^2 test. In all analyses, the pattern of results reported did not vary with different converged models. All models retained all fixed effects.

All models were built using LME models implemented in the lme4 package (Bates & Maechler, 2009) in R statistical software (R Development Core Team, 2011). For each model, I report the coefficients and standard errors for the fixed effects, and the likelihood that each coefficient differs from zero. *P*-values for continuous dependent variables were calculated by setting an upper limit for the degrees of freedom based on the number of observations minus the number of fixed effects parameters (Baayen, 2008).

All predictors were centred prior to analysis, and coded using effects coding. This procedure helps to minimise collinearity (Baayen, 2008) and means that significance tests in the mixed-effects model correspond to tests for main effects and interactions in an ANOVA model (Cohen, Cohen, West, & Aiken, 2003).

4.4. Power analysis

The probability of detecting an effect, if the effect is indeed present, is known as statistical power; the higher the statistical power, the more likely the test is to detect an effect that actually exists. Power analysis can be used to test whether a given experimental design, with a specified number of conditions and sample size, is sufficiently high-powered to detect an effect that exists in the data; lack of power may lead to the researcher committing a Type II error (Cohen, 1988). When data are analysed using classical inferential statistics such as t-test or ANOVA, the power of a given design can be calculated based on the sample size, the effect size, and the desired significance level (see, e.g., Cohen, 1988, for examples of such analysis). However, this type of classical power analysis is not recommended for use in linear mixed effects models. Instead, researchers are advised to calculate the power using simulation of fake data (Gelman & Hill, 2007).

In the simulation approach, researchers generate numerous (e.g., 1000) sets of simulated data in which they specify the hypothesised fixed effects. The size of these fixed effects is estimated based on previous research. Each simulated dataset is then analysed using the same LME model. Power is calculated as the proportion of simulations for which the model correctly reports the specified effect as being statistically significant. For example, in order to test the power of a particular design to detect a difference of 100 ms between conditions, researchers would simulate 1000 datasets in which they have specified that the 100 ms difference exists, and then analyse these datasets. This process allows researchers to see how well their design can detect an effect which they know exists in those datasets: we would expect that, in a well-powered design, the model should be able to detect the effect on ~80% of simulations (Gelman & Hill, 2007). In this thesis, post-hoc simulation-based power analysis was conducted for Experiments 3 and 4, to check whether the experimental design had sufficient power to detect an effect (see sections 5.7.3 and 6.3.3).

5. BODY-SPECIFICITY IN ACTION SENTENCES

5.1. Overview of the chapter

Embodied approaches to language propose a tight link between the motor and language systems, in which language comprehenders perform a covert motor simulation of the actions being described, in order to achieve understanding (Barsalou, 2008, 2009; Pulvermüller, 2001). In Chapter 2, I discussed the evidence that action language comprehension involves covert simulation of the described actions (see also Fischer & Zwaan, 2008; Kiefer & Pulvermüller, 2012; Meteyard et al., 2012, for reviews). However, it is unclear from which perspective comprehenders might run such simulations. For example, comprehenders might run a simulation as though they themselves were performing the described action (embodied agent), or they might run a simulation as though they were observing somebody else perform the described action (embodied observer).

In this chapter, I report four experiments that distinguish between these two possibilities, and also test whether this perspective taking is grounded in long-term motor experience, or current motor context. I do so by testing whether people who interact with the world in a particular way (e.g., performing actions with a particular hand) show evidence of interpreting action language in line with their long-term motor experience. In Experiment 1, I provide reaction time evidence that right- and left-handed participants interpret first-person sentences describing manual actions as though they were performing the described action with their dominant hand. In Experiment 2a, I provide further evidence of body-specific interpretations of action language, and also show that this tendency is stronger in first-person than in third-person sentences. In Experiment 2b, I provide evidence suggesting that these apparent effects of body-specificity in language comprehension, may in fact be driven by hand-in-use (current motor context), rather than dominant hand (long-term motor

experience). In Experiment 3, I test for, but do not find, evidence for body-specific representations of manual action sentences in the absence of overt manual responses.

5.2. Introduction

Embodied accounts of language comprehension often assume that comprehenders represent described actions from the perspective of the person performing the action – i.e., the embodied agent’s perspective (Barsalou, 1999, 2008; Zwaan & Taylor, 2006). However, as discussed in Chapter 3, the embodied agent’s perspective is not the only perspective available. In understanding a sentence such as *I am slicing the tomato*, the comprehender could equally adopt the perspective of an embodied observer (i.e., simulate watching someone else slice a tomato). One means of distinguishing which of these possible perspectives (embodied agent versus embodied observer) is adopted, is to look at the way right- and left-handed participants interpret language about actions typically performed with the dominant hand. People typically perform unimanual actions and experience greater motor fluency with their dominant hand; however, the fact that the great majority of people in the world are right-handed, means that everyone – regardless of hand dominance – observes a majority of right-handed actions in other people. In other words, right- and left-handed participants have different motor experiences, but similar visual experiences (at least of other people’s actions). Therefore, if language comprehenders tend to adopt an embodied agent’s perspective, then right- and left-handed participants should show evidence of interpreting manual action language differently from one another. But if language comprehenders tend to adopt an embodied observer’s perspective, then right- and left-handed participants should interpret manual action language in a similar way to one another.

Initial findings from a lexical decision task during fMRI suggest that right- and left-handed comprehenders interpret isolated manual actions verbs (e.g., *throw*) as though they were performing the action with their corresponding dominant hand – in other words, by adopting an embodied agent’s perspective (Willems, Hagoort, et al., 2010). However, there

is, as yet, no behavioural evidence that language comprehenders adopt a body-specific, embodied agent's perspective; and no evidence about whether this preference for the embodied agent's perspective is present in sentences as well as in isolated verbs. Indeed, the fact that motor resonance appears to be a rather localised, short-lived phenomenon (Glenberg, Sato, Cattaneo, et al., 2008; Zwaan & Taylor, 2006), coupled with evidence that action words presented in isolation may elicit higher motor activation than action sentences (Raposo et al., 2009), raises the possibility that comprehenders may not adopt an embodied agent's perspective in entire sentences – or at least, that we would not find any evidence of this at the behavioural level. In Experiment 1 therefore, I test whether left- and right-handed participants engaged in a sentence-picture matching task, show different interpretations of manual action sentences such as *I am slicing the tomato*.

5.2.1. Self and other in action language

Motor resonance in action observation is evidenced by the fact that observed and executed actions appear to share common action representations (e.g., Buccino et al., 2001; Fadiga et al., 1995; Rizzolatti & Craighero, 2004). On the other hand, there must be some distinguishing feature between our own and other people's actions, else action attribution would be impossible (Decety & Chaminade, 2003; Jeannerod, 2006; Ruby & Decety, 2001). One such feature may be the visual orientation in which an action is presented. For example, a hand performing an action could be presented with the arms extending upwards from the bottom of the image, congruent with the location of the comprehender's arms (*internal* orientation); or with the arms extending downwards from the top of the image, congruent with the location of an observed other's arms (*external* orientation).

In recent years, several findings have emerged that suggest that motor resonance in action observation may be more pronounced when observed hand is presented in an internal, rather than external orientation. For example, Maeda, Kleiner-Fisman and Pascual-Leone (2002) measured MEPs from hand muscles as participants watched videos of manual actions

that had either an internal or an external orientation. MEPs were significantly increased when participants watched the internally orientated actions, compared with the externally orientated actions (see also Alaerts, Heremans, Swinnen, & Wenderoth, 2009). Increased motor resonance in response to internally versus externally orientated hands has also been demonstrated behaviourally in visuomotor priming (Miall et al., 2006; Vogt et al., 2003), and imitation (Jackson et al., 2006).

Generalising from action perception to language comprehension, we might therefore expect that the degree of motor resonance exhibited in action language understanding to be influenced by the orientation (internal, external) of a depicted action. Brunyé and colleagues used a sentence-picture matching task with internally or externally orientation pictures (Brunyé et al., 2009). Participants judged whether sentences describing manual actions in the first-, second-, or third-person, matched internal or externally orientated images of that action. Participants were faster to match first- and second-person sentences to internally orientated images, and third-person sentences to externally orientated images. These results have been interpreted as demonstrating that comprehenders simulate the described action from the actor's perspective (adopting an embodied agent's perspective) in first- and second-person sentences, and simulate the described action from the perspective of someone observing that action (adopting an embodied observer's perspective) on third-person sentences (see also Ditman, Brunyé, Mahoney, & Taylor, 2010).

The fact that comprehenders appear to adopt a first-person perspective on sentences such as *I am slicing the tomato* may appear counterintuitive, given that participants hearing this sentence in normal dialogue would need to interpret *I* as referring not to themselves, but to their interlocutor. In fact, in a second experiment by Brunyé et al., in which the action sentences were preceded by a short narrative context (e.g., *I am a 30-year old deli employee*), participants were now faster to match the first-person sentences to externally orientated images, as though adopting an embodied observer's perspective. Why might this be the case? In section 2.2.5 we suggested that comprehenders might respond differently to

first- and third-person sentences because the third-person sentences are unambiguously about another person. Adding a preceeding context which contradicts the comprehender's circumstances may therefore perform a disambiguating function, making it clear that the *I* in the sentence should *not* be interpreted as referring to the comprehender. Without any such disambiguating context, and in the absence of an interlocuter, comprehenders may simply attribute the pronoun to the only potential speaker in that situation – themselves. The posited link between comprehension and production systems may make this more likely (see Pickering & Garrod, 2013). In the present study, sentences will be presented with no preceeding context, thus encouraging comprehenders to adopt different perspectives on the first- versus third-person sentences. Note that there is already evidence suggesting that self-referential processing may enjoy a privileged status in memory (Kelley et al., 2002; Klein & Kihlstrom, 1986; Rogers, Kuiper & Kirker, 1977). The suggestion that self-referential language may be more tightly connected to the sensorimotor system than third-person language might be viewed as an extension of this privileged processing status.

Recall from section 2.2.4, research suggests that motor experience may play a greater role in action recognition than visual experience (e.g., Loula et al., 2005). Marzoli, Mitaritonna, Moretto, Carluccio and Tommasi (2011) investigated how action orientation (internal, external) affects the relative roles of motor and visual experience on action recognition. Left- and right-handed participants were asked to imagine a third party performing a manual action (e.g., using a toothbrush), from either a front view (consistent with an external orientation) or a back view (consistent with an internal orientation). Participants then indicated with which hand (left, right) they imagined the action being performed. Participants' motor experience should encourage right-handed participants to imagine right-handed actions, and left-handed participants to imagine left-handed actions. On the other hand, participants' visual experience should encourage all participants to imagine right-handed actions. Interestingly, the results showed that when the participants imagined actions from the back view (internal orientation), they showed a greater effect of

motor experience: left-handed participants tended to imagine left-handed actions; and right-handed participants tended to imagine right-handed actions. However, when participants imagined the action from the front view (external orientation), they showed a greater effect of visual experience: both left- and right-handed participants tended to imagine right-handed actions. Thus, the relative contributions of motor and visual experience were mediated by orientation. This finding supports the possibility that orientation could be a cue to agency attribution, with internally orientated actions being attributed to self (hence right- and left-handed participants imaging right- and left-handed actions, respectively), and externally orientated actions being attributed to other (hence both groups of participants imaging right-handed actions). In Experiment 2a, I will introduce an orientation manipulation (internal, external) that will be crossed with a linguistic cue to agency attribution, namely pronoun (first-versus third-person sentences).

5.3. *The sentence-picture matching paradigm*

Matching a sentence to a drawing or photograph is a widely used paradigm in the psycholinguistics literature. Use of similar tasks in empirical research can be traced back to at least the late 1960s, when participants were typically asked to match schematic diagrams to descriptions such as *a yellow square with a red vertical stripe* (Cohen, 1969). In these early studies, researchers tended to be interested in the mechanisms involved in matching one stimulus to another (e.g., Posner, Boies, Eichelman, & Taylor, 1969), or in people's judgments about spatial location (Chase & Clark, 1971; Seymour, 1969) rather than in linguistic representations. However, Chase and Clark soon realised the importance of the paradigm for psycholinguistics (Chase & Clark, 1972; Clark & Chase, 1972), based on the observation that *"for a sentence and picture to be compared they must be represented, ultimately, in the same mental format"* (Clark & Chase, 1972, p. 473).

Since the early 1970s, various versions of sentence-picture matching tasks have been used to investigate language processing in healthy adults (e.g., Chatterjee, Southwood, &

Basilico, 1999), children (e.g., van der Lely, 1996), and patients with various forms of language impairment (e.g., Caplan & Waters, 1997; Papagno, Curti, Rizzo, Crippa, & Colombo, 2006; Small, Kemper, & Lyons, 1997). In verification versions of these tasks, participants read a sentence and then decide whether or not a subsequently presented image matches the sentence. This type of task has been successfully used to investigate embodied cognition more widely (Stanfield & Zwaan, 2001; Zwaan et al., 2002), the link between the language and motor systems (Borghi & Riggio, 2009), and perspective taking in action language comprehension (Brunyé et al., 2009; Sato & Bergen, 2013). In forced choice versions of these tasks, participants are shown a sentence, followed by a choice of pictures, and asked to select the image that best matches the sentence. This type of task has been successfully used to investigate structural priming (Branigan, Pickering, & McLean, 2005; Raffray, Pickering, & Branigan, 2007) and perspective taking in action language comprehension (Pickering et al., 2012). In both verification and forced choice versions, accuracy and reaction time data can be recorded.

During a sentence-picture matching task, participants form a representation of the sentence, maintain this representation while forming a representation of the image, and then compare the two representations (e.g., Black, Nickels, & Byng, 1991). The key premise that makes these tasks of interest to psycholinguistics is, as Clark and Chase pointed out, the idea that the sentence and image are represented in essentially the same form: the image-based representation is assumed to reflect the sentence-based representation. Thus, data from sentence-picture matching offer an insight into language comprehension because, by analysing which image participants choose to match the sentence, or how quickly they verify an image, can inform us about the nature of participants' linguistic representations.

For example, whether right- or left-handed participants have similar (right-handed) representations of manual action sentences is unclear from looking at their language use alone. But by having both groups verify whether a right- or left-handed image matches a manual action sentence (Experiments 1 & 3), we can assess if the two groups show a similar

pattern of response latencies (implying that the groups have similar linguistic representations); or if the groups show different response latencies to the different types of image (implying that the groups have contrasting linguistic representations). Similarly, by asking people to read a manual action sentence and then select a right- or left-handed image (Experiments 2a, 2b, 4, 5, 6), we can assess if the two groups show a similar preference for one type of image (implying that the groups have similar linguistic representations), or if the groups show preferences for different types of image (implying that the groups have contrasting linguistic representations).

5.4. Experiment 1

In Experiment 1, I tested whether first-person action sentences such as *I am slicing the tomato* are interpreted from an embodied agent's perspective. I did this by testing whether right- and left-handed participants adopted body-specific representations of sentences describing manual actions typically performed with the dominant hand (e.g., slicing, throwing, writing...). In a sentence-picture matching task, participants read a sentence and then saw an image depicting that action performed with a left or a right hand. I predicted that participants would show evidence of body-specific representations, in line with previous research on isolated verbs (Willems et al., 2010) and motor imagery (Willems, Toni, Hagoort, & Casasanto, 2009). In other words, right-handed participants would interpret the action sentences as though they were performing the action with their right hand, and left-handed participants would interpret the sentences as though they were performing the action with their left hand. This pattern of results would add to evidence that in self-referential language, such as first-person sentences, comprehenders adopt an embodied agent's perspective (see Chapter 3). On the other hand, if both left- and right-handed participants interpreted the sentences as though they were performing the action with the right hand, this would undermine the Body-Specificity Hypothesis, and suggest that people were adopting an embodied observer's perspective, in line with their shared

visual experience of observing more right-handed actions in the world.

5.4.1. Method

Participants

Thirty-two native English speakers took part in the experiment in return for course credit or payment. All participants had normal or corrected to normal vision, and no history of reading difficulties. Handedness was assessed using the Edinburgh Handedness Inventory (EHI; Oldfield, 1971) once testing was complete. The EHI asks participants to select their hand preference for a list of ten manual actions (see Appendix A), and then calculates a *laterality quotient*. Scores of over 40 indicate that the participant is right handed. Scores of lower than -40 indicate that the participant is left-handed. Scores of between -40 and 40 inclusive, indicate that the participant is ambi-dexterous. Test re-test reliability on the EHI is high (Ransil & Schachter, 1994), and scores on the test correlates well with other handedness measures (Bryden, 1977) as well as self-reported handedness (Ransil & Schachter, 1994). Across all experiments in the thesis, the handedness assigned by the EHI matched the handedness self-reported by all participants.

Sixteen participants in Experiment 1 were left-handed (11 females; mean age = 23 years; mean EHI score = -69, EHI score range = -41 to -100). The remaining sixteen participants were right-handed (13 females; mean age = 21 years; mean EHI score = 80, EHI score range = 55 to 100). In all experiments, all participants remained naïve to the fact that they were being recruited on the basis of handedness. The majority of left-handed participants in Experiments 1 and 2a were recruited through the construction of a research participant database: during debriefing of unrelated psychology experiments, other researchers obtained various details about the participant (e.g., age, handedness, languages spoken), along with their consent to be contacted about future experiments. The remaining left-handed participants were recruited through friends who mentioned the experiments to

their left-handed acquaintances. All right-handed participants in Experiments 1 - 6 were recruited from the research participant database. This recruitment process, although lengthy, allowed us to be sure that none of our participants were aware that they had been recruited on the basis of handedness.

Materials and design

We prepared 32 photographs showing a right hand performing a manual action such as slicing a tomato, and used photo editing software to produce 32 mirror image photographs of the same action, performed by the left hand (see Figure 5-1 for examples, and Appendix B for the full stimuli). For each pair of left- and right-handed photographs, we constructed one experimental sentence, which described the event shown in the photograph. To create an equal number of filler trials, each sentence was also paired with a photograph describing a difference event. All sentences consisted of first-person pronoun + present tense verb + direct object (e.g., *I am slicing the tomato*). Half of the filler trials formed object-mismatch trials, where the photograph and sentence depicted the same verb but a different object (e.g. slicing the bread *versus* slicing the tomato). The other half of the filler trials formed verb-mismatch trials, where the photograph and sentence depicted the same object but a different verb (e.g. slicing the bread *versus* buttering the bread). Handedness of photographs was balanced across experimental and filler trials, and within the filler trials, across verb-mismatch and object-mismatch trials. All photographs were taken from an internal orientation, so that the hands appeared at the bottom of the photograph and might plausibly be interpreted at those of participant.

Each participant saw the left-handed version of each image twice (once on a critical trial, and once on a filler trial), and the right-handed version of each twice (once on a critical trial, and once on a filler trial). The experiment was split into 4 blocks, and each participant only saw one version of each image once per block. Each block contained 16 critical and 16 filler trials (8 object-mismatch trials and 8 verb-mismatch trials). The order of block

presentation was counterbalanced across participant handedness. Trials were pseudo-randomised such that critical and filler trials featuring left- and right-handed versions of the same photograph were not presented consecutively within, or across blocks.

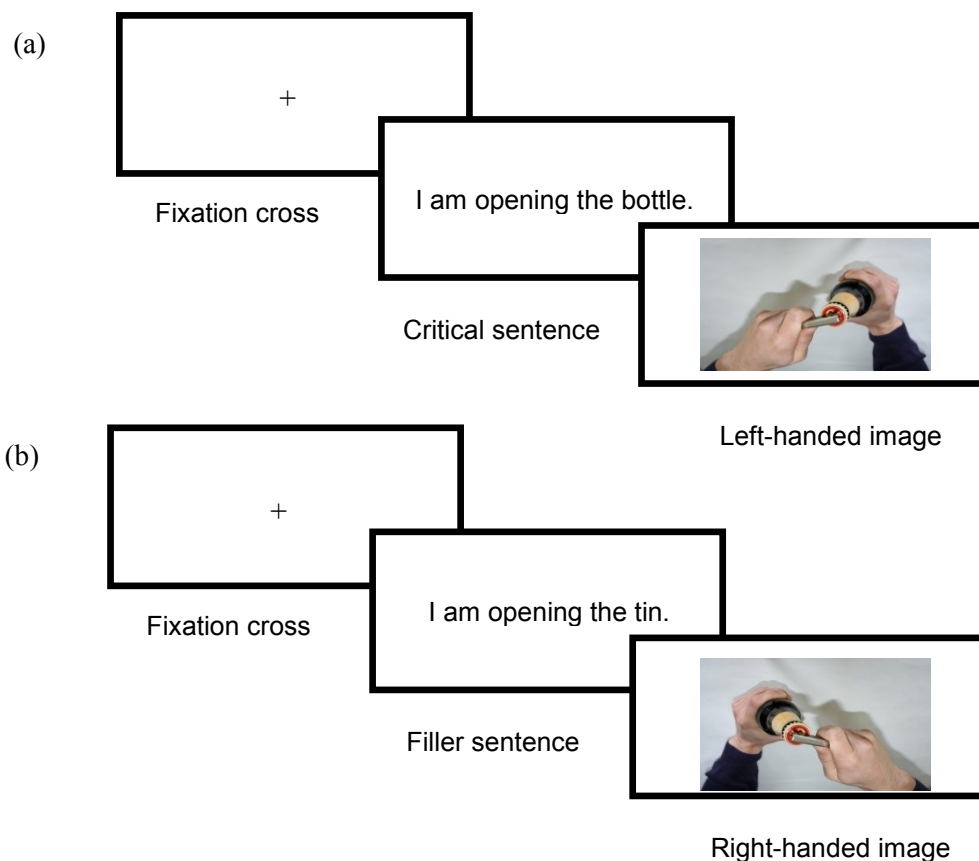


Figure 5-1 Example stimuli in critical and filler conditions in Experiment 1

Procedure

I used a sentence-picture matching task, similar to that employed by Brunyé et al., (2009). Participants sat at a computer terminal with a viewing distance of 60 cm. A button box was placed on the desk, rotated 90 degrees so that the line of buttons was perpendicular to the participant's body. Participants therefore responded by making towards and away movements rather than left and right movements, thus minimising the chances that their attention would be drawn to the left/ right handed manipulation in the images. To enable fast

responses, the button box was rotated to suit the handedness of the participant, which was noted by the experimenter as the participant filled out a consent form prior to testing.

A short practice session (6 critical trials and 6 filler trials, randomly presented) preceded the main experiment. Participants did not receive any feedback during this practice session. At the start of each trial, a central fixation cross appeared on screen. The cross remained on screen until the participant pressed the middle button on the button box with their index finger. A sentence appeared in the centre of the screen. Participants were instructed to hold down the middle button while silently reading the sentence. All text was presented in 24 point black courier new font, on a white background. After 1000 ms, the sentence was replaced by a left- or right-handed photograph depicting an action that either matched (critical trial) or did not match (filler trial) the action described in the sentence. The participant released the central button, and indicated whether the sentence and photograph matched by pressing either a nearer or a further button with the same index finger. If there was no response within 3000 ms, the trial timed out and the next trial began.

Whether participants pressed the nearer or further button to indicate “match” or “mismatch” was counterbalanced across left- and right-handed participants, and across block presentation order. I recorded RTs for (a) releasing the central button, and (b) pressing the appropriate response button. There was an enforced 90 second pause between each block. There was no mention of handedness prior to testing. All participants performed the task using their dominant hand without being prompted by the experimenter. Participants completed the EHI once testing was complete. During debriefing, participants answered the following questions: (a) What do you think was the purpose of the study? (b) On a scale of 1-10, how difficult did you find this study? (c) Did you adopt any strategy that made it easier to complete the study? (d) Looking back, can you think of any strategy that might have made the study easier? No participant mentioned dominant hand or referred to the left- and right-handedness of the images. The experiment took approximately 20 minutes, with a further 5 - 10 minutes for debriefing.

5.4.2. Results

Analysis

I analysed button release RTs and button press RTs from correct critical (match) trials, and error rates for critical trials. Predictors of interest were participant's dominant hand (left-handed participant, right-handed participant), and the hand performing the action in the image (left-handed image, right-handed image). An interaction between dominant hand and image hand would constitute evidence of body-specificity in manual action sentences, in accordance with the Body-Specificity Hypothesis (Casasanto, 2009), and suggest that participants were adopting an embodied agent's perspective on the sentences. I also checked for an effect of which direction participants were required to move to signal a match response, by including response direction (MATCH = near button, MATCH = far button) as a predictor. Unless otherwise specified, all analyses reported below use the following model with maximal random effect structure:

Model 5-1: Dominant hand * Image handedness + Response direction + (1 + Image handedness | Subject) + (1 + Participant handedness * Image handedness + Response direction | Item)

Error rates

Overall accuracy on critical trials was 85%. There were no significant differences in accuracy for left-handed participants (86%) and right-handed participants (84%), or between left-handed image choice (86%) and right-handed image choice (84%), and the interaction between these two variables was not significant (all $ps > .05$).

Button release RTs

I removed all button release RTs below 100 ms (0.5% of correct critical trials), and winsorised the remaining data so that RTs above or below 2.5 SD for participant's mean

response latency were replaced with the upper or lower cutoff value for that participant (0.1% trials replaced with lower cutoff; 2.1% trials replaced with upper cutoff). Table 5-1 shows the mean button release and button press times by condition. Table 5-2 shows the model coefficients for button release RTs using Model 5-1. As can be seen from Table 5-2, there was no main effect of dominant hand, image hand, or response direction. However, the interaction between dominant hand and image hand was significant: participants were slower to match sentences to pictures that were congruent with their own dominant hand (congruent trials = 443 (169) ms), and faster to match sentences to pictures that were incongruent with their own dominant hand (incongruent trials = 433 (162) ms). These results support the Body-Specificity Hypothesis, and suggest that comprehenders adopt an embodied agent's perspective on action language. However, if we correct for the fact that both button release and button press RTs were tested, the interaction would no longer be significant ($p > .025$).

Table 5-1. Mean winsorised button release and button press RTs (ms) by condition in Experiment 1 (sds in parentheses).

| | Left-handed image | Right-handed image |
|--------------------------|-------------------|--------------------|
| BUTTON RELEASE | | |
| Left-handed participant | 446 (184) | 435 (175) |
| Right-handed participant | 431 (149) | 440 (154) |
| BUTTON PRESS | | |
| Left-handed participant | 726 (245) | 739 (261) |
| Right-handed participant | 733 (258) | 763 (254) |

Button press RTs

I removed all responses below 200 ms (0% of correct critical trials), and winsorised the remaining data so that RTs above or below 2.5 SD for participant's mean response latency were replaced with the upper or lower cutoff value for that participant (0% trials replaced with lower cutoff; 2.0% trials replaced with upper cutoff). Table 5-1 shows the mean button release and button press times by condition. Table 5-2 shows the model coefficients button press RTs using Model 5-1. As can be seen from Table 5-2, there was no main effect of dominant hand, image hand, or response direction. The interaction between dominant hand and image hand was not significant (all $ps > .05$). The effect of body-specificity therefore appears to be limited to button release RTs.

Table 5-2. Predictors of button release and button press RTs in Experiment 1: Coefficients from Model 5-1 (significant effects shown in bold)

| Predictor | Coefficient | Standard Error | p (coefficient = 0) |
|-----------------------------------|--------------|----------------|-----------------------|
| BUTTON RELEASE | | | |
| (Intercept) | 497.2 | 27.0 | <.001 *** |
| Dominant hand | 16.2 | 29.9 | 0.9 |
| Image hand | -0.4 | 7.1 | 0.9 |
| Response direction | 21.2 | 29.5 | 0.7 |
| Dominant hand x image hand | 24.3 | 12.1 | 0.03 * |
| BUTTON PRESS | | | |
| (Intercept) | 739.8 | 28.3 | <.001 *** |
| Dominant hand | -0.3 | 54.2 | 0.9 |
| Image hand | -19.1 | 13.4 | 0.2 |
| Response direction | 18.8 | 54.2 | 0.7 |
| Dominant hand x image hand | 11.5 | 22.3 | 0.6 |

Responses to filler trials were significantly slower than critical trials for both button release RTs (critical trials = 439 (167) ms; filler trials = 479 (175) ms; $B = 36.03$, $SE = 9.01$, $p < .001$) and button press RTs (critical trials = 738 (249) ms; filler trials = 820 (251) ms; $B = 85.40$, $SE = 16.58$, $p < .001$). Within filler trials, there was no interaction between participant handedness and image handedness ($B = 9.25$, $SE = 15.82$, $p = 0.56$).

5.4.3. Discussion

In Experiment 1, I observed a significant interaction between dominant hand and image handedness in button release RTs in a sentence-picture matching task: participants were slower to verify sentence-matching for pictures involving a hand action that was congruent with their own dominant hand than for pictures that were incongruent with their own dominant hand. This interaction is in line with fMRI research suggesting that left- and right-handed participants might interpret action language according to their experience of performing actions in the world (Willems, Hagoort, et al., 2010). More generally, these results imply that comprehenders adopt the perspective of an embodied agent when understanding first-person action language, in line with the predictions from Chapter 3.

However, these findings should be treated with caution, for two reasons. First, the interaction between dominant hand and image handedness occurs for button release RTs, and not for button press RTs. The findings from button release RTs are in line with work on the ACE reporting significant effects based on button release RTs (Glenberg & Kaschak, 2002; Glenberg, Sato, Cattaneo, et al., 2008), although these studies do not report if button press RTs were also significant. Other work investigating the temporal dynamics of the ACE does report significant findings based on button press RTs (Borreggine & Kaschak, 2006; Kaschak & Borreggine, 2008), but does not report if button release RTs are also significant. When correcting for multiple comparisons (the fact that both button release and button press RTs were tested), the interaction is no longer significant. I opted to test both press and

release RTs because the literature reports roughly equal number of studies finding effects using these two dependent variables, making predicting the point at which an effect will occur (release or press), extremely difficult. It is to be hoped that future studies will report findings for both button release and button press RTs, with adjusted p-values, so that the true prevalence of effects in button releases and button presses, and the different circumstances under which these occur, can be determined. Such action might then allow researchers to better predict at what point in the response they are most likely to find an effect, and to design their studies accordingly.

Second, although I found an incongruency advantage (participants were slower to validate pictures that were congruent with their dominant hand), I did not predict an interaction in this direction, rather than a congruency advantage, given the inconsistency of effect direction in the literature. We saw in Chapter 2 that some studies investigating the link between action and language reported faster responses in congruent conditions (e.g., Borghi & Riggio, 2009; Boulenger et al., 2008; Glenberg & Kaschak, 2002; Rueschemeyer, van Rooij, Lindemann, Willems, & Bekkering, 2010; van Elk & Blanke, 2011; Zwaan & Taylor, 2006), and that other studies reported faster responses in incongruent conditions (Buccino et al., 2005; Dalla Volta, Gianelli, Campione, & Gentilucci, 2009; de Vega, Moreno, & Castillo, 2013; Mirabella, Iaconelli, Spadacenta, Federico, & Gallese, 2012; Nazir et al., 2008; Scorolli et al., 2009), and furthermore that the field is still without an adequate explanation for these differences in effect direction. Part of the reason for this lack of explanation is that, until recently, many studies aimed simply to show some evidence of some form of embodiment, and to discredit approaches that allowed no role for embodiment (see section 2.3.2). In other words, researchers tended to lapse into a dichotomy between embodiment and non-embodiment and, since an interaction of any kind was taken as evidence for the former, there was no systematic effort by the community as a whole to explore why the direction of effects should differ so strikingly between studies.

A hallmark of scientific theories is that, as they develop, they allow scientists to begin predicting not only the presence of an effect, but the type of effect that would occur. It appears that no such robust predictions are presently available to accounts of embodied cognition; working to develop them should be an urgent priority over the coming years. Until then, the reason for an incongruency advantage (rather than congruency advantage) on button release RTs (rather than button press RTs) in Experiment 1 remains unclear.

5.5. Experiment 2a

In Experiment 2a, I extended Experiment 1 by testing whether body-specificity in action language was affected by sentence pronoun (*I am* / *He is*), and image orientation (internal / external). The Spatial Grounding Hypothesis (see Chapter 3) predicts that action-perspective taking will be reduced in non-self-referential language (e.g., third-person sentences), compared with self-referential language (e.g., first-person sentences). If this is correct, then participants should be less likely to adopt an embodied agent's perspective for third-person sentences – in other words, body-specificity should be stronger for first-person sentences than for third-person sentences. Research suggests that participants may show stronger sensorimotor activation in response to internally orientated hands (Marzoli et al., 2011; Vogt et al., 2003). Therefore, participants might also be less likely to adopt an embodied agent's perspective on externally orientated images than on internally orientated images (see Figure 5 – 2 for examples of internally and externally orientated stimuli).

5.5.1. Method

Participants

Thirty-two native English speakers took part in the experiment in return for course credit or payment. All participants had normal or corrected to normal vision, and no history of reading difficulties. Handedness was assessed using the Edinburgh Handedness Inventory

(EHI; Oldfield, 1971) once testing was complete. Sixteen participants were left-handed (10 females; mean age = 25 years; mean EHI score = -80.10, EHI score range = -60 to -100). Sixteen participants were right-handed (13 females; mean age = 20 years; mean EHI score = 92, EHI score range = 66 to 100).

Materials and design

Materials were the same as for Experiment 1, with the addition of third-person versions (*he is...*) of each critical and filler sentence; and externally orientated versions of each right- and left-handed photograph (see Appendix B). Because Experiment 2 used a two-alternative forced choice matching task, each sentence was now paired with 2 images (one right-handed, one left-handed). On critical trials, both images depicted the action described in the sentence. On filler trials, one image depicted the action described in the sentence, and one image depicted a different action that did not match the sentence. In all trials, both images in a pair were the same orientation (either both internal or both external). In filler items, the image handedness was counterbalanced across matching and mismatching images.

The design of Experiment 2 was as follows. Subject pronoun (*I am / He is*), and image orientation (internal / external) were fully crossed, giving four within-participants and within-items conditions:

First-person + Internal

Third-person + Internal

First-person + External

Third-person + External

The experiment consisted of four blocks, each with 32 critical and 32 filler items (16 object-mismatch trials and 16 verb-mismatch trials), so that each participant saw four items per condition in each block, and each item appeared once per condition (in a different condition per block) for each participant. The order of blocks was counterbalanced across participants in a Latin square design. In order to control for an effect of spatial congruency (right-handed picture on the right of the screen versus the left of the screen), I also balanced,

across participants and items, which image appeared on the left or right side of the screen. Therefore, for each participant, half the trials in each condition were spatially congruent, and half were spatially incongruent. Each item appeared in a spatially congruent trial for half the participants, and a spatially incongruent trial for the remaining participants. The order of trials was randomised within block.

Procedure

I used a forced-choice sentence-picture matching task, similar to that employed by Pickering et al. (2012). Apparatus and computer set up were the same as Experiment 1, with the exception that, due to the setup of the dual button box apparatus, the button boxes were placed horizontally facing the participant rather than at a 90 degree angle. A short practice session (4 critical trials, 12 filler trials) took place before the main experiment. Participants did not receive any feedback during this practice session. At the start of each trial, a central fixation cross appeared. The cross remained on screen until the participant pressed the middle button on a button response box with their index finger. A sentence appeared in the centre of the screen. Participants were instructed to hold down the middle button while silently reading the sentence. After 1000 ms, the sentence was replaced by a pair of images, one left-handed and one-right handed. One of the images appeared on the left of the screen, and the other image appeared on the right of the screen. The participant selected one of the two photographs to match the preceding sentence, by releasing the middle button on the button box and pressing either the left or right button with the same index finger, to select the picture on the left side or right side of the screen, respectively. On critical trials, both photographs showed the action described in the sentence, but with a different hand (left or right) performing the action in each image. If no response was made within 3000 ms, the trial timed out. In addition to image choice, I recorded RTs for (a) releasing the central button, and (b) selecting an image.

There was an enforced 90 second pause between each of the 4 blocks. On 25% of trials (4 randomly selected critical trials and 12 randomly selected filler trials per block) the word REPEAT appeared on screen following image selection, to cue participants to repeat the sentence out loud. This repetition task was introduced in order to disguise the purpose of the experiment, and to ensure that participants attended to the pronoun as well as the action. Participants' oral responses were recorded and coded according to whether participants recalled the sentence correctly or not.

There was no mention of handedness or "left" and "right" prior to the experiment. The experimenter explained how the participant was to respond using the terms "this side of the screen" and "this button", rather than "left side of the screen" or "right button". Participants' handedness was assessed by observing the participant fill out a consent form prior to the experiment, and by completion of the EHI during debriefing following the experiment. All participants performed the task using their dominant hand without being prompted by the experimenter.

5.5.2. Results

Analysis

I analysed the likelihood of selecting a right-handed image on critical trials, and error rates for response choice on filler trials, and for sentence repetition on all trials. In addition, I analysed the button release and button press RTs from correct critical trials. Predictors of interest were dominant hand (left-handed participant, right-handed participant), pronoun (first-person, third-person), and image orientation (internal, external). A main effect of dominant hand would add to the evidence from Experiment 1 for body-specificity in manual action sentences, and suggest that participants were adopting an embodied agent's perspective on the sentences. An interaction between dominant hand and pronoun would suggest that the extent to which participants adopt an embodied agent's perspective is

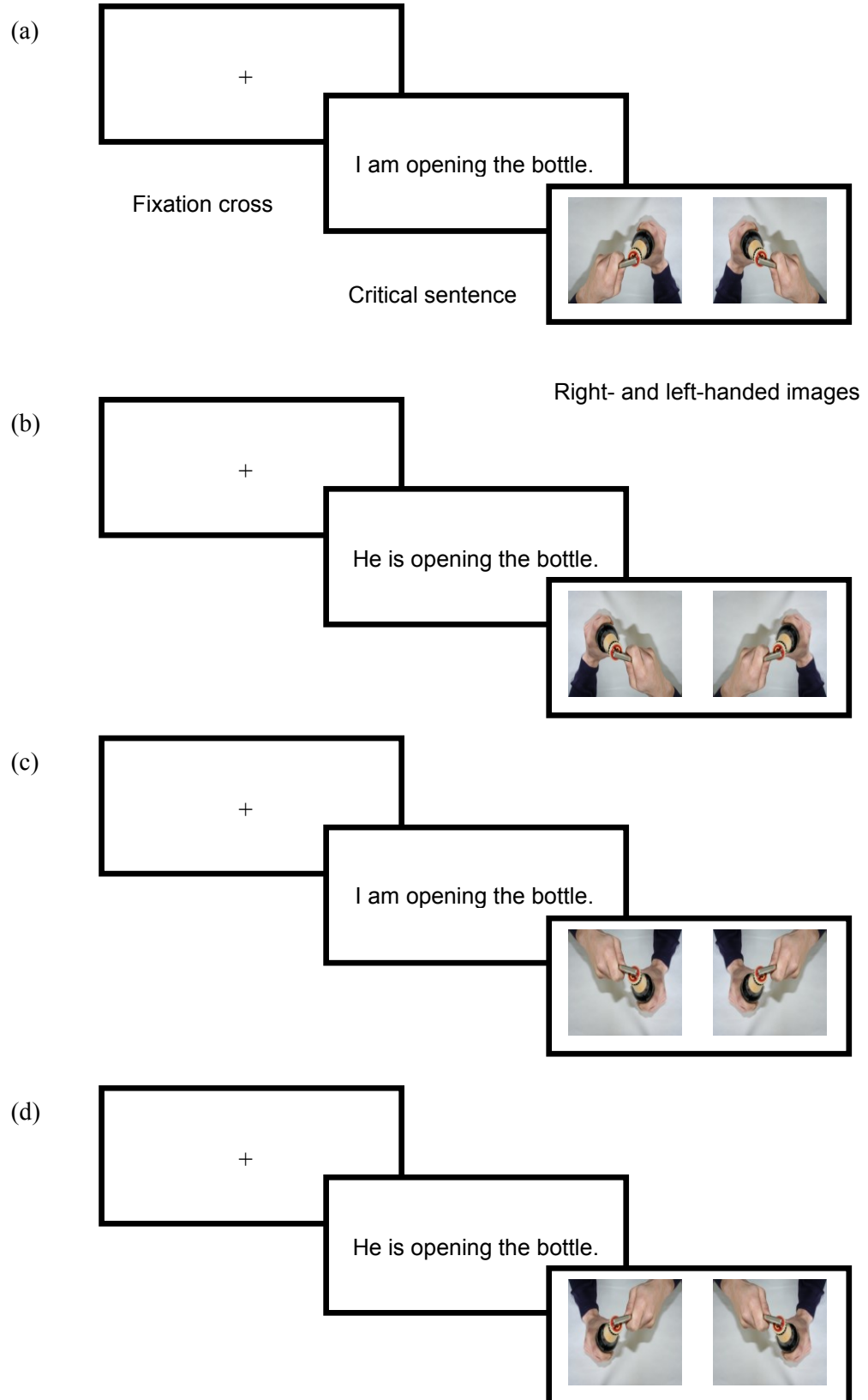


Figure 5-2. An experimental item in the four conditions in Experiment 2a

modulated by the linguistic context. An effect of image orientation would suggest that participants used orientation as a cue to agency attribution, and that their interpretation of action language differed according to this cue. I also expected a spatial-compatibility effect where participants would be more likely to select a right-handed image when it was positioned on the right side of the screen (cf. Hommel, 2011; Rubichi & Nicoletti, 2006), and therefore included image-screen congruency (CONGRUENT = right-handed image on right side of screen, INCONGRUENT = right-handed image on left side of screen) as a predictor.

Unless otherwise specified, all analyses reported below use the following model with simplified random effect structure (the correlation parameter, and the random intercepts and slopes for image-screen congruency, were removed in order to allow the model to converge):

Model 5-2: Dominant hand * Pronoun * Image orientation + Image-screen congruency + (1 | Subject) + (1 | Item) + (0 + Pronoun * Image orientation | Subject) + (0 + Dominant hand * Pronoun * Image orientation | Item)

Error rates

Sentence-picture matching

Overall accuracy on selecting the correct image to match the sentence on filler trials was 94%. There were no significant differences in accuracy for first-person sentences (94%) versus third-person sentences (93%); for internally (93%) versus externally (94%) orientated images; or for image-screen congruent (94%) versus incongruent (95%) trials (all $ps > .05$). Right-handed participants were significantly more accurate (96%) than left-handed participants (93%) in the task ($B = 0.74$, $SE = 0.21$, $p < .001$).

Sentence repetition

Overall accuracy for sentence repetition was 94% (97% on critical trials; 92% on filler trials, *ns.*). There were no significant differences in accuracy for first-person sentences (94%) versus third-person sentences (94%); for internally (95%) versus externally (93%) orientated images; for image-screen congruent (95%) versus incongruent (93%) trials; or for right-handed (95%) versus left-handed (93%) participants (all $ps > .05$).

Image choice

In total, 6.8% of critical trials timed-out without a response. I removed the 0.02% trials where participants selected an image within 200 ms of image onset. Tables 5-3 and 5-4 show the frequency of image choice by condition. Table 5-5 shows the model coefficients for the likelihood of selecting a right-handed image. There was a main effect of dominant hand: right-handed participants were more likely than left-handed participants to select a right-handed image. In addition, this preference interacted with pronoun: participants were more likely to select an image that was congruent with their own hand dominance following a first-person sentence, compared with a third-person sentence. There was also a significant effect of image-screen congruency (see Table 5-3). There were no effects of image orientation, as a main effect, or as an interaction (all $ps > .05$).

Table 5-3. Frequency of right-handed and left-handed image choice by image-screen congruency in Experiment 2a (percentage responses in parentheses).

| Dominant hand | Image-screen congruency | Image choice | |
|---------------|--|-------------------|--------------------|
| | | Left-handed image | Right-handed image |
| Left hand | <i>Right-handed image on right of screen</i> | 764 (76%) | 236 (24%) |
| | <i>Left-handed image on right of screen</i> | 525 (56%) | 412 (44%) |
| Right hand | <i>Right-handed image on right of screen</i> | 399 (40%) | 576 (60%) |
| | <i>Left-handed image on right of screen</i> | 415 (45%) | 491 (55%) |

Table 5-4. Frequency of right-handed and left-handed image choice by handedness, pronoun, and image orientation in Experiment 2a (percentage responses in parentheses)

| | Pronoun | Image orientation | Left-handed image | Right-handed image |
|---------------------------|---------------------------------------|--------------------------|--------------------------|---------------------------|
| Left-handed participants | I am | Internal | 333 | 148 |
| | I am | External | 333 | 158 |
| | Total I am (all orientations) | | 666 (69%) | 306 (31%) |
| | He is | Internal | 308 | 178 |
| | He is | External | 315 | 164 |
| | Total He is (all orientations) | | 623 (65%) | 342 (35%) |
| Right-handed participants | I am | Internal | 197 | 280 |
| | I am | External | 183 | 283 |
| | Total I am (all orientations) | | 380 (40%) | 563 (60%) |
| | He is | Internal | 220 | 250 |
| | He is | External | 214 | 254 |
| | Total He is (all orientations) | | 434 (46%) | 504 (54%) |

Button Release RTs

I removed trials with button release RTs of under 100 ms (3.3% critical trials), and winsorised remaining critical trials so that RTs above or below 2.5 SD for participant's mean response latency were replaced with the upper or lower cutoff value for that participant (0.01% trials replaced with lower cutoff; 2.3% trials replaced with upper cutoff). Tables 5-6 and 5-7 show the mean button release RTs, by condition. I tested for effects of dominant hand, pronoun, image orientation, and image-screen congruency using Model 5-2 (all $ps > .05$).

Table 5-5. Predictors of image choice in Experiment 2a: Coefficients from Model 5-2 (significant fixed effects in bold)

| Predictor | Coefficient | Standard Error | <i>p</i> (coefficient = 0) |
|---|-------------|----------------|----------------------------|
| <i>Likelihood of choosing right-handed image</i> | | | |
| (Intercept) | -0.29 | 0.14 | <.05 * |
| Dominant hand | 1.17 | 0.28 | <.001 *** |
| Pronoun | 0.02 | 0.08 | 0.84 |
| Image orientation | 0.05 | 0.16 | 0.76 |
| Image-screen congruency | 0.28 | 0.09 | <.01 ** |
| Dominant hand x pronoun | 0.50 | 0.15 | <.001 *** |
| Dominant hand x image orientation | 0.09 | 0.35 | 0.79 |
| Pronoun x image orientation | 0.08 | 0.15 | 0.58 |
| Dominant hand x pronoun x image orientation | 0.13 | 0.30 | 0.68 |

Button press RTs

I removed trials with button press RTs of under 200 ms (0.02% critical trials), and winsorised remaining critical trials so that RTs above or below 2.5 SD for participant's mean response latency were replaced with the upper or lower cutoff value for that participant (0.04% trials replaced with lower cutoff; 2.14% trials replaced with upper cutoff). Tables 5-6 and 5-7 show the mean button press RTs, by condition. I tested for effects of pronoun, image orientation, dominant hand, and image-screen congruency using Model 5-2 (all *ps* > .05).

Table 5-6. Mean winsorised button release and button press RTs (ms) by image-screen congruency in Experiment 2a (sd in parenthesis)

| Dominant hand | Image-screen congruency | Image choice | |
|----------------|---------------------------------------|-------------------|--------------------|
| | | Left-handed image | Right-handed image |
| BUTTON RELEASE | | | |
| Left hand | Right-handed image on right of screen | 624 (304) | 564 (294) |
| | Left-handed image on right of screen | 701 (336) | 543 (311) |
| Right hand | Right-handed image on right of screen | 576 (319) | 580 (348) |
| | Left-handed image on right of screen | 577 (312) | 597 (359) |
| BUTTON PRESS | | | |
| Left hand | Right-handed image on right of screen | 1023 (403) | 975 (385) |
| | Left-handed image on right of screen | 1112 (398) | 882 (373) |
| Right hand | Right-handed image on right of screen | 1123 (570) | 1172 (555) |
| | Left-handed image on right of screen | 1104 (496) | 1155 (534) |

Table 5-7. Mean winsorised button release and button press RTs (ms) by response hand, pronoun, and image orientation in Experiment 2a (sd in parentheses)

| | Left-handed image choice | | | | Right-handed image choice | | | |
|---------------------------|--------------------------|------------|------------|------------|---------------------------|------------|------------|------------|
| Pronoun | I am | He is | I am | He is | I am | He is | I am | He is |
| Image orientation | Internal | Internal | External | External | Internal | Internal | External | External |
| BUTTON RELEASE | | | | | | | | |
| Left-handed participants | 639 (283) | 660 (285) | 661 (336) | 661 (365) | 541 (255) | 540 (286) | 545 (331) | 576 (328) |
| Right-handed participants | 569 (272) | 583 (331) | 598 (346) | 558 (308) | 591 (331) | 554 (304) | 585 (377) | 618 (389) |
| BUTTON PRESS | | | | | | | | |
| Left-handed participants | 1017 (391) | 1057 (387) | 1095 (420) | 1056 (434) | 996 (359) | 1022 (432) | 1060 (365) | 1003 (474) |
| Right-handed participants | 1031 (440) | 1147 (547) | 1104 (562) | 1156 (600) | 1136 (548) | 1070 (483) | 1194 (600) | 1181 (585) |

5.5.3. Discussion

In Experiment 2a, I replicated an effect of body-specificity on how comprehenders interpret action language sentences: in a two-alternative forced-choice sentence-picture matching task, right-handed participants were more likely to select a right-handed image, and left-handed participants were more likely to select a left-handed image. These results reinforce the findings from Experiment 1, which appear to show that people adopt an embodied agent's perspective, rather than an embodied observer's perspective in action language comprehension.

Importantly, the effect of dominant hand interacted with pronoun such that right-handed participants were more likely to select a right-handed image following a first-person sentence compared with a third-person sentence, and left-handed participants were more likely to select a left-handed image following a first-person sentence compared with a third-person sentence. This interaction demonstrates that the effect cannot solely be due to differences between right- and left-handers' processing of right- and left-handed images. Work on affordances has shown that participants are faster to respond to objects that are displayed as optimally "graspable" for the hand making the response (Tucker & Ellis, 1998). Our results from Experiment 1 might therefore have been interpreted in terms of the left- and right-handed images being optimally graspable by left- and right-handed participants, respectively. However, the interaction with sentence pronoun in Experiment 2a indicates that there cannot *only* be an effect of image-based affordances. At the very least, these image-based affordances are interacting with language processing.

Interestingly, there was no effect of image orientation. This was slightly unexpected, given work suggesting that motor resonance is stronger in response to internal rather than external images (Vogt et al., 2003), that the role of handedness interacts with image orientation (Marzoli et al., 2011), and that image orientation and subject pronoun interact with one another in action language comprehension (Brunyé et al., 2009).

5.6. Experiment 2b

In Experiment 1a, all participants responded using their dominant hand (without any prompting from the experimenter), and appeared to adopt an embodied agent's perspective on sentences describing manual actions. One explanation of these results is an interaction between language and long-term motor experience, consistent with the Body-Specificity Hypothesis (Casasanto, 2011), whereby a comprehender's understanding of action language is grounded in their long-term motor experience of the world (i.e., the experience of having used their dominant hand to perform the actions described in the sentences). However, the results are also compatible with an interaction between language and current motor context (i.e., the hand being used to respond in this particular task). Experiment 2b was therefore designed to dissociate response hand from dominant hand.

In Experiment 2b, I replicated Experiment 2a with a group of right-handed participants, but this time participants were instructed to respond using their non-dominant (i.e., left) hand. In this way, we aimed to test whether the results from Experiments 1 and 2a were due to long-term motor experience (i.e., the dominant hand), or the immediate motor context of the task (i.e., the hand-in-use). I also wished to determine whether the effect of pronoun, and null effect of image orientation replicated. If the apparent effects of body-specificity in Experiments 1 and 2a are due to long-term motor experience, then right-handed participants responding with their left hand should behave like the right-handed participants in Experiment 2a. If the apparent effects of body-specificity in Experiments 1 and 2a are in fact due to current motor context, then right-handed participants responding with their left hand should behave like the left-handed participants in Experiment 2a.

5.6.1. Method

Participants

Sixteen native English speakers (9 females, mean age = 22.4 years, age range = 18 - 29) took part in return for course credit or payment. All participants had normal or corrected to normal vision, and no history of reading difficulties. All participants were right-handed according to the EHI (mean EHI score = 85, EHI score range = 60 to 100).

Materials and design

Materials and design were the same as Experiment 2a.

Procedure

The procedure was the same as Experiment 2a, except that participants responded to all trials using their non-dominant left hand. Participants were informed they were taking part in a study on working memory, and would be using their left hand to make the task more difficult. No other mention of handedness was made prior to testing, and no participant noticed a connection between dominant or response hand and handedness of the images.

5.6.2. Results

Analysis

Analysis was the same as Experiment 2a, with the following exception. Trial number within the experiment (i.e., the point in the experiment at which a particular response was made) was now included as a fixed effect in all models, and as a random effect when the resulting model converged. This inclusion aimed to account for the fact that participants may become more fluent with their left hand as the task progressed. Unless otherwise specified in the text, all analyses reported below used the following model with simplified random effect structure (the correlation parameter, and the random intercepts and

slopes for image-screen congruency, were removed in order to allow the model to converge):

Model 5-3: Pronoun * Image orientation + Image-screen congruency + Trial + (1 |Subject) + (1| Item) +(0+ Pronoun * Image orientation + Trial | Subject) + (0+ Pronoun * Image orientation | Item)

Error rates

Sentence-picture matching

Overall accuracy on selecting the correct image to match the sentence on filler trials was 95%. There were no significant differences in accuracy for first-person sentences (95%) versus third-person sentences (95%); for internally (95%) versus externally (95%) orientated images; or for image-screen congruent (96%) versus incongruent (94%) trials (all $ps > .05$).

Sentence repetition

Overall accuracy for sentence repetition was 96% (97% on critical trials; 95% on filler trials, *ns.*). There were no significant differences in accuracy for first-person sentences (94%) versus third-person sentences (96%); for internally (97%) versus externally (95%) orientated images; or for image-screen congruent (96%) versus incongruent (96%) trials (all $ps > .05$).

Image choice

In total, 7% of critical trials timed-out without a response and could not be analysed. I removed any further trials where participants selected an image within 200 ms of image onset (1.37% remaining critical trials). Tables 5-8 and 5-9 show the frequency with which participants selected a right-handed image, by condition. Table 5-10 shows the model coefficients for the likelihood of selecting a right-handed image using Model 5-3. We can see from Table 5-9 that right-handed participants using their left hands chose left-handed images more often than right-handed images. Moreover, there was a significant effect of

pronoun: right-handed participants using their left hand were more likely to select a left handed image following first-person sentences compared with third-person sentences. There was also a main effect of image-screen congruency, with participants more likely to choose a left-handed image when it appeared on the left side of the screen (68%) compared with the right of the screen (46%). There was no effect of image orientation, and no effect of trial: participants were equally likely to select a right-handed image throughout the experiment.

Button release RTs

I removed trials with button release RTs of under 100 ms (0.02% critical trials), and winsorised remaining critical trials so that RTs above or below 2.5 SD for participant's mean response latency were replaced with the upper or lower cutoff value for that participant (0% trials replaced with lower cutoff; 2.69% trials replaced with upper cutoff). Tables 5-11 and 5-12 show the mean button release RTs, by condition. I analysed the button release RTs using Model 5-3. There were no significant effects of pronoun, image orientation, or image-screen congruency, and no significant interactions between these variables. There was no effect of trial on button release RTs (all $ps > .05$).

Button press RTs

I removed trials with button press RTs of under 200 ms (1.37% critical trials), and winsorised remaining critical trials so that RTs above or below 2.5 SD for participant's mean response latency were replaced with the upper or lower cutoff value for that participant (0.05% trials replaced with lower cutoff; 2.17% trials replaced with upper cutoff). Tables 5-11 and 5-12 show the mean button press RTs, by condition. I analysed the button press RTs using Model 5-3. There were no significant effects of pronoun, image orientation, or image-screen congruency, and no significant interactions between these variables (all $ps > .05$). There was no effect of trial on button release RTs (all $ps > .05$).

Table 5-8. Frequency of right-handed and left-handed image choice by image-screen congruency in Experiment 2b (percentage responses in parentheses)

| Response hand | Image-screen congruency | Image choice | |
|---------------|--|-------------------|--------------------|
| | | Left-handed image | Right-handed image |
| Left hand | <i>Right-handed image on right of screen</i> | 676 (68%) | 319 (32%) |
| | <i>Left-handed image on right of screen</i> | 413 (46%) | 494 (54%) |

Table 5-9. Frequency of right-handed and left-handed image choice by pronoun and image orientation in Experiment 2b (percentage responses in parentheses)

| Pronoun | Image orientation | Left-handed image | Right-handed image |
|---------------------------------------|-------------------|-------------------|--------------------|
| I am | Internal | 302 | 182 |
| I am | External | 278 | 193 |
| Total I am (all orientations) | | 580 (61%) | 375 (39%) |
| He is | Internal | 255 | 214 |
| He is | External | 254 | 224 |
| Total He is (all orientations) | | 479 (52%) | 438 (48%) |

Table 5-10. Predictors of image choice in Experiment 2b: Coefficients from Model 5-3 (significant fixed effects in bold)

| Predictor | Coefficient | Standard Error | p (coefficient = 0) |
|---|-------------|----------------|---------------------|
| <i>Likelihood of choosing right-handed image</i> | | | |
| (Intercept) | -0.37 | 0.16 | 0.02* |
| Pronoun | 0.31 | 0.13 | 0.02* |
| Image orientation | 0.08 | 0.12 | 0.50 |
| Image-screen congruency | 0.98 | 0.11 | <.001*** |
| Pronoun x image orientation | -0.12 | 0.20 | 0.57 |

Table 5-11. Mean winsorised button release and button press RTs (ms) by image-screen congruency in Experiment 2b (sd in parenthesis)

| Response hand | Image-screen congruency | Image choice | |
|-----------------------|--|-------------------|--------------------|
| | | Left-handed image | Right-handed image |
| BUTTON RELEASE | | | |
| Left hand | <i>Right-handed image on right of screen</i> | 618 (306) | 660 (326) |
| | <i>Left-handed image on right of screen</i> | 594 (330) | 591 (335) |
| BUTTON PRESS | | | |
| Left hand | <i>Right-handed image on right of screen</i> | 1045 (443) | 1104 (427) |
| | <i>Left-handed image on right of screen</i> | 1104 (497) | 1023 (474) |

Between experiment comparisons

In order to confirm whether right-handed participants using their left hand performed significantly differently from left-handed participants using their left hand, I conducted a planned between experiment comparison in which I compared the choice data of right-handed participants in Experiment 2b with that of left-handed participants in Experiment 2a, using the following model with simplified random effect structure (the random intercepts and slopes for image-screen congruency and the dominant hand by trial interaction term were removed in order to allow the model to converge):

Model 5-4: Dominant hand * Pronoun * Image orientation + Image-screen congruency + Dominant hand x Trial + (1+ Pronoun * Image orientation | Subject) + (1+ Dominant hand * Pronoun * Image orientation + | Item)

Table 5-12. Mean winsorised button release and button press RTs (ms) by condition in Experiment 2b (sd in parentheses)

| | Left-handed image choice | | | | Right-handed image choice | | | |
|---------------------------------|--------------------------|------------|------------|------------|---------------------------|------------|------------|------------|
| <i>Pronoun</i> | I am | He is | I am | He is | I am | He is | I am | He is |
| <i>Image orientation</i> | Internal | Internal | External | External | Internal | Internal | External | External |
| <i>BUTTON RELEASE</i> | 634 (329) | 640 (355) | 645 (395) | 663 (420) | 576 (329) | 557 (308) | 577 (332) | 573 (319) |
| <i>BUTTON PRESS</i> | 1041 (423) | 1058 (455) | 1095 (482) | 1086 (487) | 1027 (457) | 1029 (467) | 1042 (509) | 1057 (536) |

Table 5-13 shows the model coefficients for this between-experiments comparison. Left-handed participants were more likely to select a left-handed image than were right-handed participants using their left hand, but importantly, this trend did not reach significance. In addition, there was a main effect of pronoun: participants were significantly more likely to choose a left-handed image following a first-person sentence than following a third-person sentence. There was also a significant effect of image-screen congruency, with participants more likely to choose a left-handed image when it appeared on the left of the screen than on the right of the screen. There was no effect of image orientation, nor any further interaction between image orientation, pronoun and dominant hand (all $ps > .05$). The interaction between trial and dominant hand was not significant.

In order to confirm whether there were any significant differences between right-handers using their left hand, and right-handers using their right hand, I conducted a second planned between-experiment comparison. Here, I compared the image choices of right-handed participants in Experiment 2b (responding with left hand) with those of right-handed participants in Experiment 2a (responding with right hand). Results showed a main effect of response hand: participants tended to select an image that was congruent with their response hand. There was no main effect of pronoun, but there was a significant interaction between response hand and pronoun: people were more likely to select an image congruent with their response hand following first-person sentences, compared with third person sentences. There was a significant effect of image-screen congruency: participants were more likely to select a right-handed image when it appeared on the right of the screen. There was no effect of image orientation, trial, nor any other interaction between image orientation, pronoun and response hand (all $ps > .05$).

Table 5-13. Predictors of image choice in Experiments 2a and 2b: Coefficients from Model 5-4 (significant fixed effects in bold)

| Predictor | Coefficient | Standard Error | p (coefficient = 0) |
|---|--------------|----------------|---------------------|
| <i>Likelihood of choosing right-handed image</i> | | | |
| Left-handers Expt. 2a / Right-handers Expt. 2b | | | |
| (Intercept) | -0.62 | 0.15 | <.001*** |
| Dominant hand | 0.44 | 0.32 | 0.17 |
| Pronoun | 0.32 | 0.10 | 0.001*** |
| Image orientation | 0.13 | 0.12 | 0.30 |
| Image-screen congruency | 1.03 | 0.10 | <.001*** |
| Dominant hand x Pronoun | 0.01 | 0.19 | 0.95 |
| Dominant hand x Image orientation | -0.08 | 0.19 | 0.69 |
| Pronoun x Image orientation | -0.06 | 0.17 | 0.74 |
| Dominant hand x Trial | 0.0005 | 0.0007 | 0.42 |
| Dominant hand x Pronoun x Image orientation | -0.09 | 0.21 | 0.77 |
| Right-handers Expt. 2a / Right-handers Expt. 2b | | | |
| (Intercept) | -0.03 | 0.10 | <.80 |
| Response hand | 0.62 | 0.16 | <.001*** |
| Pronoun | 0.01 | 0.19 | .87 |
| Image orientation | 0.09 | 0.13 | .47 |
| Image-screen congruency | 0.42 | 0.08 | <.001*** |
| Response hand x Pronoun | -0.56 | 0.15 | <.001*** |
| Response hand x Image orientation | -0.006 | 0.24 | 0.98 |
| Pronoun x Image orientation | -0.07 | 0.14 | .61 |
| Response hand x Trial | -0.02 | 0.03 | .53 |
| Response hand x Pronoun x Image orientation | -0.08 | 0.27 | .76 |

5.6.3. Discussion

Right-handed participants who responded using their left hand in Experiment 2b showed a preference for left-handed images on critical trials. This pattern of results parallels that of the left-handed participants using their left-hand in Experiment 2a, and contrasts with that of the right-handed participants using their right hand in Experiment 2a. The absence of any effect of trial number in Experiment 2b implies that participants do not simply “become” more left-handed as the experiment progresses. These data suggest that comprehenders adopt an embodied agent’s perspective on manual action sentences, and that this perspective is based on the hand being used to respond to the task, rather than the hand that would typically be used to perform the described actions.

In Experiment 2b, I found a main effect of pronoun: for right-handed participants using their left hand, the preference for left-handed images was stronger following first-person sentences, compared with third-person sentences. This effect of pronoun is in line with the interaction observed between pronoun and dominant hand in Experiment 2a, suggesting that the extent to which participants adopt an embodied agent’s perspective can be modulated by linguistic context, and the degree to which the language used to describe the actions can be interpreted as referring to the comprehender herself. Once again, we observed no effect of image orientation, either as a main itself or in interaction with other predictor variables.

These data suggest that when participants respond with their non-dominant hand, this overrides the long-term motor experience of their dominant hand. However, it may be the case that in the absence of any manual response, the effect of long-term motor experience becomes apparent. For example, Willems et al., (2010) found evidence suggesting that dominant hand affects action language representations when neither hand is required to make a response. Comparing Experiments 2a and 2b, we can see that although the right-handed participants in Experiment 2b preferred left-handed images in general, their

preference was weaker than that of the left-handers in Experiment 2a, although this trend was not significant. Experiment 3 was designed to test for an effect of dominant hand in the absence of manual responses.

5.7. Experiment 3

In Experiment 2b, right-handed participants using their left hand were more likely to select left-handed images than right-handed images, and this preference was stronger following first-person than in third-person sentences. These results suggest that comprehenders adopt an embodied agent's perspective based on hand-in-use (current motor context), rather than dominant hand (long-term motor experience). In other words, we have found an effect of current motor context. However, an underlying effect of long-term motor experience remains a possibility. In other words, the results of Experiment 2b might represent a case where the comprehender's default source of motor activation (long-term motor experience) is overridden by a conflicting motor context. To investigate this possibility, I conducted a further experiment in which there was no conflicting current motor context that might override such an effect. Experiment 3 was designed to test for an effect of body-specificity in a vocalized response task in which no overt hand actions were required.

In Experiment 3, I used a verification sentence-picture matching task as in Experiment 1. I used both first- and third-person sentences to check for an effect of pronoun, as had been observed in Experiments 2a and 2b. If comprehenders typically adopt an embodied agent's perspective driven by long-term motor experience, until this is overridden by conflicting motor context, then right-handed participants responding without moving or planning to move their hands should show different latencies in response to left- and right-handed images, in line with their long-term motor experience. In particular, given that Experiment 1 showed an incongruency advantage, we would expect right-handed participants to respond more slowly to right-handed images than to left-handed images.

5.7.1. Method

Participants

Twenty four native English speakers (10 females, mean age = 22.8 years, age range = 19 - 32) took part in return for course credit or payment. All participants had normal or corrected to normal vision, and no history of reading difficulties. All participants were right-handed according to the EHI (mean EHI score = 77, EHI score range = 41 to 100).

Materials and design

Materials and design were the same as Experiment 1, with the exception that we produced third-person versions of each critical and filler sentence, thus doubling the length of the experiment. An item consisted of a sentence (first or third person) paired with an internally orientated image (right- or left-handed). As with Experiment 1, in critical trials, the sentence and image depicted the same event; in filler trials, the sentence and image depicted different events. The experiment was divided into 4 blocks, with a different version of each item (*I am* + right-handed image; *I am* + left-handed image; *He is* + right-handed image; *He is* + left-handed image) appearing once per block.

Procedure

The procedure was the same as Experiment 1, except that participants responded to all trials vocally, instead of using their hand. Participants responded “good” if the sentence and picture matched, and “bad” if the sentence and picture did not match. If no response was made within 3000 ms of the image appearing on screen, the trial timed out and the participant moved onto the next trial. Voice Onset Times (VOTs) were recorded using the sound key on an SR response box. The experimenter recorded the participants’ response on a computer in the same room, but separated from the participant by a screen. Auditory responses were also recorded using a USB microphone to allow the experimenter to cross-check responses. There was a 90 second break between each block. The experiment was

preceded by a short practice block of 24 trials. No mention of handedness was made prior to testing, and no participant noticed a connection between dominant hand and the handedness of the images.

5.7.2. Results

Analysis

I analysed VOTs from correct critical (match) trials, and error rates between critical and filler trials, and within critical trials. Predictors of interest were pronoun (first-person, third-person), and image handedness (left-handed image, right-handed image). A main effect of image handedness of VOTs (participants being slower to respond to right-handed images) would suggest that, in the absence of conflicting immediate motor context, comprehenders interpret manual action sentences according to how they would normally perform that actions being described (in this case, with their right hand). An interaction between image handedness and pronoun would support the finding from Experiments 2a and 2b, that the embodied agent's perspective is modulated by linguistic context. Unless otherwise specified in the text, all analyses in this section are based on the following model with maximal random effect structure:

Model 5-5: Image handedness * Pronoun + (1 + Image handedness* Pronoun| Subject) + (1 + Image handedness* Pronoun | Item)

Error rates

Overall accuracy on the task was 89%. On critical trials, participants were significantly more likely to correctly match sentences to right-handed images (94% accuracy) than to left-handed images (92% accuracy; $B = 0.46$, $SE = 0.16$, $p = .004$). There was no significant difference in accuracy between first-person sentences (92% accuracy) and

third-person accuracy (94% accuracy), and no significant interaction between image handedness and pronoun (all $ps > .05$).

Voice Onset Times

Of the critical trials, I removed trials where participants corrected their response (0.8%), where the participants' response was preceded by a disfluency (1%), and where the participant repeated her response (0.04%; total trials excluded = 1.84% of critical trials). In addition, 1.60 % of all critical trials timed out without any VOT response, and were excluded from analyses. I then checked for responses made faster than 350 ms, and then winsorised the data so that VOTs above or below 2.5 SD for participant's mean VOT latency were replaced with the upper or lower cutoff value for that participant (0% trials replaced with lower cutoff; 2.56% trials replaced with upper cutoff). Table 5-14 shows the mean VOTs by condition. Table 5-15 shows the model coefficients for VOTs using Model 5-5. There was no main effect of image handedness or pronoun, and the interaction between the two was not significant (all $ps > .05$).

Table 5-14. Mean winsorised Voice Onset Times (ms) by condition in Experiment 3 (sd in parentheses)

| Pronoun | Left-handed image | Right-handed image |
|----------------|--------------------------|---------------------------|
| <i>I am</i> | 698 (212) | 693 (198) |
| <i>He is</i> | 693 (207) | 694 (202) |

Table 5-15. Predictors of Voice Onset Times in Experiment 3: Coefficients from Model 5-5

| Predictor | Coefficient | Standard Error | p (coefficient = 0) |
|--------------------------------|-------------|----------------|-----------------------|
| <i>Voice Onset Time</i> | | | |
| (Intercept) | 696.00 | 35.86 | <.001*** |
| Pronoun | -0.06 | 10.17 | 0.97 |
| Image hand | -10.46 | 26.67 | 0.18 |
| Pronoun x image hand | 3.52 | 23.28 | 0.64 |

5.7.3. Discussion

In Experiment 3, right-handed participants who responded without using their hands showed no sign of body-specificity in action sentences. There was no effect of participants being slower to respond to right-handed images, as in Experiment 1, and no effect of pronoun as had been observed in Experiment 2a and 2b. Therefore, these data do not provide any evidence that, in the absence of overt manual responses, comprehenders adopt an embodied agent's perspective based on their long-term motor experience.

Given that Experiment 1 only revealed a difference of ~15 ms, a post-hoc power analysis was conducted to check that the design in Experiment 3 had given sufficient power to detect a small effect size. This power analysis was conducted using the simulation-based approach outlined in section 4.4. I generated 1000 datasets, based on the design of Experiment 3. The fixed effect of image handedness was estimated based on the results from Experiment 1, namely, participants (all right-handed) should be 15ms slower to respond to right-handed images than to left-handed images ($b = 15$). The fixed effects for pronoun and the interaction between image handedness and pronoun were set at $b = 0$, since the estimated coefficient of these effects does not impact on the likelihood that a simulation can detect an effect in the fixed effect of interest (image handedness). The variances associated with the

random effects were estimated based on the variances in the actual data from Experiment 3, namely 16108 for subjects, 1182 for items, and 39971 for residuals. I then tested how effectively Model 5-5 (the maximal model, used in analysis of our actual data), could correctly detect the 15 ms effect of image handedness that had been specified in the simulated data. A significant difference was detected in only 21% of simulations, indicating that the experimental design was insufficiently powered to reliably detect an existing effect in image handedness. The null result in Experiment 3 may thus have been the result of a Type II error. I therefore conducted a second power analysis, using the following model which collapsed across pronoun:

Model 5-6: Image handedness + (1 + Image handedness| Subject) + (1 + Image handedness| Item)

Again, the fixed effect for image handedness was set at $b = 15$, but fixed effects of pronoun and the interaction were not specified, as the analysis collapsed over pronoun. Using this simplified model, the 15 ms difference between left- and right-handed images was correctly detected on 78% simulations, indicating reasonable power to detect an effect of image handedness. The actual data obtained from Experiment 3 were then reanalysed using Model 5-6 to check for any effect of image handedness when collapsing across pronoun. Results of this additional analysis showed no effect of image handedness on VOT ($b = -1.33$, $SE = 13.10$, $p = 0.90$), despite the increase in power.

5.8. General discussion

In four experiments, I explored evidence that comprehenders adopt an embodied agent's perspective when understanding action sentences – that is, they understand the sentences as though they were performing the action themselves. In Experiments 1 and 2a, I found evidence for body-specificity in action sentences from two different sentence-picture

matching tasks: participants were slower to correctly match sentences to images that were congruent with their dominant hand, and were more likely to select an image that was congruent with their dominant hand in a two-alternative forced choice task. These results are in line with the suggestion that when understanding language about actions, people adopt an embodied agent's perspective based on the way they have previously performed those actions. In Experiment 2b however, I found evidence that the apparent body-specificity in Experiments 1 and 2a may have been due to the current motor context rather than long-term motor experience. Right-handed participants who responded using their left hand were more likely to select a left-handed image to match a sentence in a two-alternative forced-choice task. In Experiment 3, I failed to find any evidence that in the absence of overt manual response, comprehenders interpret action language in light of their long-term motor experience: right-handed participants showed no difference in Voice Onset Times to right-handed and left-handed images in a sentence-picture matching task.

Taken together, these results suggest firstly, that comprehenders tend to adopt an embodied agent's perspective, especially for first-person language, and secondly, that this perspective is grounded in the current motor context rather than long-term motor experience. The absence of any effects of trial when right-handed participants used their left hand to respond suggests that our results cannot easily be attributed to a purely practice effect.

5.8.1. Sentence pronoun

In Experiments 2a and 2b, the hand-in-use preference was stronger following first-person compared with third-person sentences. The Spatial Grounding Hypothesis (see Chapter 3) predicts differences between self-referential (e.g., first-person) and non-self-referential (e.g., third-person) language. According to the Spatial Grounding Hypothesis, in order to run an action simulation (and thus, adopt an embodied perspective), comprehenders must be able to ground that simulation in a sufficiently detailed representation of the space in which that action would be performed. In other words, a comprehender must have access

to some coarse-coded information about the spatial properties of the action. For example, in order to simulate the action of closing a drawer, the comprehender must first be able to conceive the direction of such an action. In the case of self-referential language, comprehenders can ground themselves in their own body: a sentence such as *I am closing the drawer* can be understood in terms of moving the hand towards the comprehender's own body. But in the case of third-person language, external cues might also be required to enable the comprehender to know in which direction they should simulate a movement (e.g., spatial avatars; see Gianelli et al. , 2011). In the absence of such external cues, running an action simulation of third-person language should be more difficult compared with first-person language.

The effects of pronoun of Experiments 2a and 2b support the claims of the Spatial Grounding Hypothesis to some extent, since participants appear more likely to adopt an embodied agent's perspective for first-person, compared with third-person sentences. However, note that left-handed participants, in particular, still displayed a strong preference (~65%) for left-handed images following third-person sentences. One possible explanation of this finding is that, just as Taylor and Zwaan argue that comprehension is not an all or nothing process (Taylor & Zwaan, 2009, 2013), running a simulation during language comprehension is not all or nothing either. Action simulations may degrade gracefully in clarity or specificity as the spatial context becomes less than optimal (e.g., in third-person language). For example, motor representations of one's own actions are thought to be more detailed than representations of other people's actions (Jeannerod, 2006), a mechanism which may allow us to distinguish between our own and other's actions. Therefore, third-person language may simply trigger less detailed simulations, and therefore result in weaker embodiment effects than first-person language.

A second possibility is that third-person language does not result in less detailed simulations, but is simply less likely to result in a simulation happening at all. The idea of graceful decay (inherent in the proposal that comprehension is fault tolerant) lead us to

favour the first possibility, that third-person representations are less embodied, rather than the second possibility, that third-person representations are less likely to be embodied at all. However, the present data are insufficient to distinguish between these two possibilities empirically. Note that the results of our right-handed participants in Experiment 2a do however suggest that third-person language is not more likely to trigger an embodied observer's perspective (*contra* Brunyé et al., 2009): for right-handed participants responding using their right hand, adopting an embodied observer's perspective should also manifest itself in a preference for right-handed images.

5.8.2. Image orientation

Perhaps surprisingly, we found no effect of image orientation in any of the three groups of participants in Experiments 2a and 2b, either as a main effect or in interaction with other predictors. These null effects contrast with findings from action observation (Vogt et al., 2003), motor imagery (Marzoli et al., 2011), and action language comprehension (Brunyé et al., 2009; Pickering et al., 2012), all of which find effects of orientation (internal, external) on processes through to involve motor resonance. Why, then, do we find no effect of orientation?

Gardner and Potts (2010) asked participants to make speeded laterality judgements about whether a ball was in the left or right hand of a schematic human figure. The human figure was shown from either a front view (external orientation) or a back view (internal orientation). Left-handed participants were faster to respond when the ball was in the figure's left hand, and right-handed participants were faster to respond when the ball was in the figure's right-hand, regardless of internal or external orientation. Gardner and Potts hypothesise that attentional biases towards an individual's dominant hand (e.g., Amazeen, Amazeen, Treffner, & Turvey, 1997; Amazeen, Ringenbach, & Amazeen, 2005; Rubichi & Nicoletti, 2006) generalise from an individual's own body to other, observed bodies (or in our case, hands). However, this explanation stresses the primacy of dominant hand

representations, both behaviourally (e.g., Summers, Semjen, Carson, & Thomas, 1995) and cortically (Triggs, Subramaniam, & Rossi, 1999). Therefore, although it can account for the findings of Experiment 2a, it cannot easily account for the reversal of behaviour in Experiment 2b, where right-handed participants showed a preference for left-handed images.

Instead, it seems that hand orientation only inconsistently acts as a cue to agency attribution. For example, Anquetil and Jeannerod (2007) found that participants were equally as fast to imagine a manual action from an external orientation as from an internal orientation. The conditions in which orientation does and does not act as a cue are unclear, and should be a topic for further research. However, a clue may lie in a recent TMS study, where participants observed manual actions in which either the muscle performing the movement, or the direction that movement, could be congruent with the participant (Alaerts, Swinnen, & Wenderoth, 2009). Results showed that MEPs were modulated by the congruency of the muscle rather than that of the direction. Extrapolating to our study, it may be that congruency of the hand performing the action, rather than its orientation, took precedence. Notably, in those studies which found an effect of orientation, only one study (Marzoli et al., 2011) manipulated the hand dominance of participants, thus manipulating the effector as well as the direction; and in Marzoli et al., only a single trial was used per participant.

5.8.3. *Current motor context*

The results of Experiment 2b suggest that although participants appear to adopt an embodied agent's perspective, this is driven by current motor context rather than long-term experience. In one sense then, the results support the idea of body-specificity, since an individual's current actions appear to influence their linguistic representations. However in another sense, the results conflict with the idea of body-specificity, since we found no evidence that a comprehender's history of interaction in the world influences their language comprehension. Previous work on body-specificity has found that training can reverse the

typically observed preferences (Casasanto & Chrysikou, 2011), however, trial showed no effect when it was included as an effect in Experiment 2b. In other words, people did not appear to become “more left-handed” as the experiment progressed. We will explore these issues further in Chapter 6.

For now however, it is interesting to note that an influence of current motor context would fit well with the construal of embodiment type effects as a form of priming (section 2.3.5): the current state of motor activation influences the current state of linguistic processing, but is not necessary or sufficient for language comprehension to occur (to say otherwise would be to imply that participants in Experiment 3 did not understand the sentences they were presented with).

5.8.4. Demand characteristics

One potential worry with the studies in this chapter is that participants may have guessed our intention to investigate differences in handedness (Experiment 1 and 2a), or response hand (Experiment 2b). Cues that unintentionally convey the experimental hypothesis to the participant are known as demand characteristics (Orne, 1962). Participants can use these demand characteristics to adjust their behaviour in order to meet what they believe to be the expectations of the experimenter, in particular for tasks requiring participants to make a subjective decision about, for example, mental imagery (Intons-Peterson, 1983). Given the subjective nature of the task in Experiments 2a and 2b (there is no objectively correct image on the critical trials), we are therefore justified in worrying about the potential effects of demand characteristics. I sought to rule out the effects of demand characteristics as much as possible in the following four ways. First, the repetition of sentences on one quarter of trials provided a plausible cover story to disguise the true purpose of the experiment. Second, the recruitment of participants was designed in such a way that no participant realised they had been recruited on the basis of their handedness. Third, image selection was made under time pressure, with the average

response being initiated at ~610 ms. Fourth, a thorough debriefing after the experiment allowed us to check whether any participant had guessed at the handedness manipulation. One participant in Experiment 2a mentioned that he thought we were testing whether he would select the right-handed or the left-handed image, and he was removed from all further analysis. All suitable precautions were therefore taken to reduce the possibility of demand characteristics affecting the results reported here.

5.9. Summary

In this chapter, I found that language comprehenders appear to adopt an embodied agent's perspective on action language: that is, interpreting it as if they themselves were performing the action. This embodied agent's perspective interacts with subject pronoun, so that participants' preference for an interpretation that matches their current state of motor activation is stronger following first-person sentences compared with third-person sentences. I also found evidence that the embodied agent's perspective is driven by current motor context rather than previous experience. I found no evidence for an effect of image orientation (internal versus external).

6. DEFINING THE CURRENT MOTOR CONTEXT

6.1. Overview of the chapter

In Chapter 5, I found evidence that people interpret action language sentences in line with hand-in-use rather than dominant hand. Participants were more likely to match manual action sentences to images that were congruent with participants' response hand, rather than with their dominant hand. This preference for images depicting the hand-in-use was stronger following first-person sentences compared with third-person sentences. In other words, it appears that current motor context overrides long-term motor experience in the interpretation of action language.

However, *current motor context* is rather vaguely defined. For example, the effects described in Chapter 5 were observed in image selection, after sentence processing had been completed. It is therefore unclear whether an effect of current motor context entails that motor resonance be present during sentence processing (sentence-based simulation), or whether motor resonance and its interaction with pronoun might arise after sentence processing is complete (image-based simulation). It is also unclear what the basis of this motor resonance might be: in Experiments 2a and 2b, hand-in-use refers to both the hand the participants were planning to respond with on a given trial, and the hand that participants had responded with on the previous trial. In this chapter, I extend the forced choice sentence-picture matching paradigm used in Chapter 5 so that participants have one response box for their left hand, and one response box for their right hand.

In this way, I aim to specify the parameters of *current motor context* at the following two levels. First, I manipulate at what point in time (before sentence processing, after sentence processing), participants are cued as to which hand to respond with on a given trial. Second, I manipulate whether response hand is randomised or blocked, in order to dissociate response hand on the current trial from response hand on the previous trial. In Experiment 4,

I provide evidence that the effect of hand-in-use and its interaction with pronoun disappear when participants do not know their respond hand until after sentence processing. In Experiments 5 and 6, I show that when participants know their respond hand prior to sentence processing, the effects of hand-in-use and pronoun reappear. Experiment 6 suggests that the difference between Experiments 4 and 5 is not an artefact of blocking (Experiment 5) versus randomisation (Experiment 4), and furthermore, that current motor context is defined by the hand comprehenders are planning to use (prospective simulation), rather than the hand used to respond on a previous trial (retrospective simulation).

6.2. Introduction

The results from Chapter 5 suggest that comprehenders interpret manual action sentences according to their hand-in-use rather than their dominant hand, and that this tendency is stronger following self-referential language, in line with the Spatial Grounding Hypothesis proposed in Chapter 3. The interaction found between hand-in-use and pronoun is important because that it rules out the possibility that the effect of hand-in-use might simply be an effect of left- and right-hand based affordances in image processing (cf. Tucker & Ellis, 1998), with no interaction with language: at some point prior to image selection, the likelihood of selecting an image congruent with hand-in-use is affected by whether the action is described in first- or third-person language. Two factors remain unclear, however. First, *when* does this interaction between the language and action systems occur? Second, does hand-in-use have an effect because it represents the hand that participants are planning to use, or because it represents the hand used to respond on the previous trial?

Regarding the first question, one possibility is that the language system represents first- and third-person sentences differently, but that these representations are not sensitive to motor activation from hand-in-use. That is, there is no hand-in-use based action simulation during sentence processing. Motor resonance instead occurs during image

processing (perhaps as a result of left- or right-hand based affordances), and this activation interacts with the different linguistic representations for first- and third-person language after sentence processing is complete. In other words, the effect of pronoun occurs during sentence processing, but the effect of hand-in-use, and its interaction with pronoun, occurs *after* sentence processing. A second possibility is that the language system is sensitive to motor activation from hand-in-use, and simulations based on this activation occur during sentence processing. Thus, comprehenders tend to represent the described action as if carried out by their response hand. As discussed in section 5.8.1, this motor representation is either stronger, or more likely to occur, for first-person compared with third-person language, perhaps as a result of motor representations of our own actions being stronger than those of other people (Jeannerod, 2006). In other words, the effects of pronoun and of hand-in-use, and their interaction, all occur *during* sentence processing. Experiments 4 and 5 will aim to distinguish between these two possibilities.

Regarding the second question, there is evidence to suggest that short-term motor experience can affect our representations of valence and affect (Casasanto & Chrysikou, 2011), and of motor resonance in action observation (Petroni et al., 2010), supporting the idea that current motor context is defined by the hand used to perform previous responses. However, other evidence suggests that motor resonance is greater for actions we predict compared with those we do not predict (Kilner, Vargas, Duval, Blakemore, & Sirigu, 2004), supporting the idea that current motor context is defined by hand used to plan for future responses. Experiment 6 aims to shed light on this question by randomizing response hand so that the hand participants plan to respond with can be dissociated from their response hand on the previous trial.

6.2.1. Temporal dynamics and motor resonance

At what point does the interaction between hand-in-use and pronoun occur? Above, I sketched two possibilities: that the effect of hand-in-use occurs after sentence processing,

and interacts differentially with first-person and third-person interpretations; or that the effect of hand-in-use occurs during sentence processing, and is stronger for first-person compared with third-person sentences. In each of these explanations, the language system (pronoun) interacts with motor activation (hand-in-use). However, only in the second explanation (hand-in-use and pronoun interact *during* sentence processing), is it assumed that motor resonance is present during sentence comprehension. Therefore, only the second explanation implies a functional role of motor resonance.

Much of the debate about the possible functional role of motor resonance in language processing revolves around at what point in language processing motor activation can be detected. For example, evidence that the motor system is activated by 250 ms (or earlier) after word onset, has been found using fMRI (Pulvermüller, Härle, & Hummel, 2001), MEG (Pulvermüller, Shtyrov, & Ilmoniemi, 2005), and in behavioural studies (Boulenger et al., 2006). These studies are often cited as evidence that embodied representations play a functional role in semantic meaning, rather than being a downstream by-product of the comprehension process (e.g., Pulvermüller & Fadiga, 2010). Of course, evidence that motor resonance occurs early on in language processing is still compatible with there being no causal relationship between the two systems, since both processes may simply begin at the same time without interacting with one another (Mahon & Caramazza, 2008). Therefore, demonstrating that the effect of hand-in-use occurs during sentence processing is not sufficient grounds for inferring a *causal* relationship between the action and language systems. However, evidence that motor resonance only occurs after language processing has been completed, is *not* compatible with the motor system playing a functional role in language comprehension.

One means of discerning whether the effect of hand-in-use occurs during or after sentence processing is to manipulate at what point in time participants know which hand they will use to respond with. If motor activation from hand-in-use interacts with pronoun during sentence processing, then the effects of hand-in-use and pronoun should only occur

when participants know their response hand during sentence processing. If motor activation from hand-in-use interacts with pronoun after sentence processing, then the effects of hand-in-use and pronoun should occur also when participants do not know their response hand until after sentence processing is complete. In Experiments 4 – 6 therefore, I manipulate whether participants know their response hand before, or after sentence processing. This manipulation draws on work looking at the time course of the Action-Sentence Compatibility Effect, where researchers manipulate at what point in processing participants know which in direction they should respond (towards the body or away from the body) to a sensible sentence (Borreggine & Kaschak, 2006; Diefenbach et al., 2013; Kaschak & Borreggine, 2008).

Recall from Chapter 2 that timing has been demonstrated to affect the presence, or the direction of effects in action language research. Importantly for the present chapter, Borreggine and Kaschak (2006) found an ACE when participants were told the response-direction mapping at sentence onset, but not when this mapping was presented at various time points (50ms, 500ms, 1000ms) after the end of the sentence. According to the authors, the ACE occurs only when participants are given the chance to plan their motor response *during* sentence processing.

6.2.2. Motor planning versus short-term motor experience

Throughout Chapter 5, I described the effect of hand-in-use as an effect of *current motor context*. However, in Experiments 2a and 2b, participants responded with the same hand throughout the experiment. Therefore, the motor context could be defined in one of two ways. First, it could be that the crucial factor is the hand that was used to respond on the previous trial. This possibility would suggest that comprehenders make use of retrospective simulation, where motor resonance is based on simulation of actions performed in the (recent) past. Alternatively, it could be that the crucial factor is the hand that the participant is planning to use on the current trial. This possibility would suggest that comprehenders

make use of prospective simulation, where motor resonance is based on simulation of upcoming actions. Both of these possibilities have some support from the literature, and Experiment 6 will test between them.

In favour of retrospective simulation, we know that training people to experience increased motor fluency in their non-dominant hand causes them to associate positive valence with their non-dominant hand rather than their dominant hand (Casasanto & Chrysikou, 2011). Glenberg, Sato and Cattaneo (2008) demonstrated that plasticity in embodied representations can affect language comprehension: participants spent approximately 20 minutes moving beans either away from their body or towards their body, before performing sensibility judgements on sentences depicting movements away from or towards the body. Results showed that participants were slower to perform sensibility judgements on sentences describing movement that was congruent with the movement in their training period. The null effect of trial over the course of Experiment 2b suggests that participants in that experiment did not gradually become more “left-handed” as the experiment progressed. However, it is possible that simulations might have been based on residual activation from the previous response (and that this activation did not build up cumulatively through the experiment, hence participants were not more “left-handed” at the end of the experiment than at the beginning). In this case, a single previous trial might be sufficient to determine to motor context.

On the other hand, Rueschemeyer and Bekkering (2013) point to the fact that motor resonance in action observation is often viewed as having a predictive role (Blakemore & Frith, 2005; Wilson & Knoblich, 2005). For example, participants were shown video clips of a hand, and cued at the start of each clip as to whether or not the hand in that clip would move. On the cued-movement clips, EEG revealed motor activation *prior* to the observation of any actual movement (Kilner et al., 2004). Evidence of anticipatory motor resonance has been demonstrated in infants as young as nine months old. Southgate et al. (2009) applied EEG to infants as they performed reaching actions, and as they observed others reaching for

an object. On observation trials, the object was visible for 400ms before a hand appeared and performed the reaching movement. Results showed similar attenuation in alpha-band activity on observation trials as on performing trials. Interestingly, the infants showed attenuation during the first 400 ms of observation trials, (i.e., before any reaching movement was observed), but this early attenuation only occurred on later trials, when the infant had learnt that a reaching movement would soon occur. The effect of trial in this study suggests that infants did not simply respond to visible objects, but rather predicted upcoming actions regarding those objects.

Finally, still in favour of prospective simulation, van Elk, van Schie and Bekkering (2009) describe a study in which unusual actions appear to overturn the effect of more typical actions rooted in long-term experience. Adult participants were instructed to perform actions with typical goals (e.g., bringing a cup to the mouth) or atypical goals (e.g., bringing a cup to the eye) in response to words that could be congruent with the typical goal based on past experience, or with the unusual goal based on their current motor plan. For example, when moving a cup to their mouth, the word *mouth* would be congruent with both the experience and motor planning based goals. When moving a cup to their eye, the word *mouth* would be congruent with the experience-based goal, and the word *eye* would be congruent with the motor planning based goal. As expected, participants were faster to perform usual actions in response to words that were congruent with the experience-based goal. However, participants were faster to perform unusual actions in response to words that were congruent with the motor planning based goal, thus overriding the effects of previous experience. Therefore, if we conceive the task of responding with a non-dominant hand as an unusual action goal, it may be the case that the motor activation associated with this atypical action plan overrides the activation associated with long-term dominant hand experience.

6.3. Experiment 4

In Experiment 4, I conducted a within-subjects version of Experiments 2a and 2b, in which participants did not know until after sentence processing which hand they would use to respond with. I adapted the forced-choice sentence-picture matching paradigm used in the previous chapter, so that participants had a coloured response box for their left hand, and a different coloured response box for their right hand. Participants were shown a coloured screen to cue which box (and therefore hand) they should respond with, before sentence presentation. If the hand-in-use by pronoun interaction from Experiments 2a and 2b was due to sentence-based simulations, which were stronger for first-person than for third-person language, then there should be no effect of pronoun or hand-in-use on the likelihood of choosing a right-handed image in this version of the task, since motor resonance will not be available during sentence processing. Alternatively, if the hand-in-use by pronoun interaction in Experiments 2a and 2b was driven by image-based simulations interacting with different first- and third-person linguistic representations *after* sentence processing, then we should replicate the results from Experiments 2a and 2b here: a main effect of hand-in-use and an interaction between hand-in-use and pronoun.

6.3.1. Method

Participants

Sixteen native English speakers (11 females, mean age = 19.3 years, age range = 17 – 35) took part in return for course credit or payment. All participants had normal or corrected to normal vision, and no history of reading difficulties. All participants were right-handed according to the EHI (mean EHI score = 86, EHI score range: 58 to 100).

Materials and design

I employed a two-alternative forced-choice sentence-picture matching task, as in Experiments 2a and 2b. All sentences and image pairs were identical to those used in Experiments 2a and 2b. Design was the same as Experiment 2a and 2b, with the following exceptions. Participants now had two response boxes, one green coloured, and one blue coloured. The response hand (left, right) using a particular response box (blue box, green box) was counterbalanced across participants so that half the participants used their left hand to respond with the blue box, and their right hand to respond with the green box; the other half of the participants used their right hand to respond with the blue box, and their left hand to respond with the green box. Person (*I am, He is*) and image orientation (internal, external) remained fully crossed. Response hand (and hence, screen cue: blue or green) was counterbalanced across filler and critical trials, and across condition (First-person + Internal, First-person + External, Third-person + Internal, Third-person + External).

Procedure

Procedure was the same as Experiments 2a and 2b, with the following exceptions. Participants sat with the index finger of one hand resting on, but not pressing, the middle button of the blue response box, and the index finger of the other hand resting on, but not pressing, the middle button of the green response box. A fixation cross appeared on screen (1000 ms), followed by the sentence (1000 ms). Next, participants were shown a blue or green coloured screen (1500 ms). The screen colour represented the box (and therefore, the hand) that participants should use to respond with their choice of image on that trial. The coloured screen was followed by a pair of photographs, one on the left of the screen, and one on the right of the screen. Participants selected the image on the right or left of the screen by pressing the right or left button, respectively, on the response box that they had been cued to use on that trial. If a response was not made within 3000ms, the trial timed out.

The experiment consisted of 4 blocks, with an enforced 90 second pause between each block. To ensure that participants attended to the subject of the sentence as well as the action described, and to disguise the purpose of the experiment, participants were required to repeat the sentence aloud following 25% trials. Participants' oral responses were recorded and coded according to whether participants had recalled the sentence correctly or not.

Presentation of trials was randomised. There was no mention of handedness or "left" and "right" prior to the experiment. The experimenter explained how the participant was to respond using the terms "this side of the screen" and "this button", rather than "left side of the screen" or "right button". Participants' handedness was checked by observing the participant fill out a consent form prior to the experiment, and by completion of the EHI during debriefing following the experiment. A short practice session (16 trials) preceded the experimental session.

6.3.2. Results

Analysis

Image choice, sentence repetition accuracy and button press RT analyses were conducted in the same manner as Experiments 2a and 2b. Error rate analysis differed in that participants could now be incorrect on filler trials in either their image choice, or the response hand used to make that choice. On critical trials, participants could be either correct or incorrect in the response hand used to select an image.

Predictors of interest were response hand (left-handed response, right-handed response), sentence pronoun (first-person, third-person), and image orientation (internal, external). An interaction between response hand and pronoun would demonstrate that image-based motor resonance could interact with sentence pronoun, after sentence processing is complete. An effect of image orientation would conflict with the null effect of

image orientation in Experiments 2a and 2b. We also expected a spatial-compatibility effect as per Experiments 2a and 2b. Image-screen compatibility (CONGRUENT = right-handed image on right side of the screen, INCONGRUENT = right-handed image on left side of the screen) was therefore included as a fixed effect in all models. Unless otherwise specified, all analyses reported below use the following model with simplified random effect structure, where random intercepts and slopes for image-screen congruency were removed in order to allow the model to converge:

Model 6-1: Response hand * Pronoun * Image orientation + Image-screen congruency + (1+ Response hand * Pronoun * Image orientation | Subject) + (1+ Response hand * Pronoun * Image orientation | Item)

Error rates

Sentence-picture matching

Overall accuracy on selecting the correct image to match the sentence on filler trials was 94%. There were no significant differences in accuracy for first-person sentences (94%) versus third-person sentences (94%); for internally (94%) versus externally (94%) orientated images; for right-handed (93%) versus left-handed (94%) responses; or for image-screen congruent (94%) versus incongruent (93%) trials (all $ps > .05$).

Sentence repetition

Overall accuracy for sentence repetition was 97% (98% on critical trials; 96% on filler trials, ns.). There were no significant differences in accuracy for first-person sentences (97%) versus third-person sentences (96%); for internally (96%) versus externally (98%) orientated images; for image-screen congruent (95%) versus incongruent (98%) trials; or for right-handed (98%) versus left-handed (98%) participants (all $ps > .05$).

Response hand

Overall accuracy for using the correct response hand was 96% (96% on critical trials; 95% on filler trials, ns.). There were no significant differences in accuracy for first-person sentences (96%) versus third-person sentences (95%); for internally (96%) versus externally (95%) orientated images; for image-screen congruent (96%) versus incongruent (96%) trials; or for right-handed (96%) versus left-handed (95%) responses (all $ps > .05$).

Image choice

In total, 1% of critical trials timed-out without a response, and could not be analysed. I checked for any trials where participants selected an image within 200 ms of image onset (0 trials). Tables 6-1 and 6-2 show the frequency with which participants selected a right-handed image, by condition. Table 6-3 shows the model coefficients for the likelihood of selecting a right-handed image, using the following model with simplified random effect structure (the correlation parameter, the random intercepts and slopes for image-screen congruency, and the by-item random intercepts and slopes for interactions with item orientation were removed in order to allow the model to converge):

Model 6-2: Response hand * Pronoun * Image orientation + Image-screen congruency + (1+ Response hand * Pronoun * Image orientation | Subject) + (1+ Response hand * Pronoun + Image orientation | Item)

As can be seen from Table 6-3, there was no main effect of response hand, pronoun, nor any interaction between response hand and pronoun (all $ps > .05$). There was a significant effect of image-screen congruency (see Table 6-1). There were no effects of image orientation, as a main effect, or in interaction with other predictors (all $ps > .05$).

Table 6-1. Frequency of right-handed and left-handed image choice by image-screen congruency in Experiment 4 (percentage responses in parentheses).

| Response hand | Image-screen congruency | Image choice | |
|---------------|--|-------------------|--------------------|
| | | Left-handed image | Right-handed image |
| Left hand | <i>Right-handed image on right of screen</i> | 150 (31%) | 334 (69%) |
| | <i>Left-handed image on right of screen</i> | 298 (63%) | 178 (37%) |
| Right hand | <i>Right-handed image on right of screen</i> | 153 (32%) | 321 (68%) |
| | <i>Left-handed image on right of screen</i> | 335 (67%) | 167 (33%) |

Table 6-2. Frequency of right-handed and left-handed image choice by response hand, pronoun, and image orientation in Experiment 4 (percentage responses in parentheses)

| Response hand | Pronoun | Image orientation | Image choice | |
|---------------|--|-------------------|-------------------|--------------------|
| | | | Left-handed image | Right-handed image |
| Left hand | <i>I am</i> | <i>Internal</i> | 107 | 137 |
| | <i>I am</i> | <i>External</i> | 117 | 119 |
| | Total <i>I am</i> (all orientations) | | 224 (47%) | 256 (53%) |
| | <i>He is</i> | <i>Internal</i> | 110 | 136 |
| | <i>He is</i> | <i>External</i> | 114 | 120 |
| | Total <i>He is</i> (all orientations) | | 224 (47%) | 256 (53%) |
| Right hand | <i>I am</i> | <i>Internal</i> | 123 | 123 |
| | <i>I am</i> | <i>External</i> | 129 | 114 |
| | Total <i>I am</i> (all orientations) | | 252 (52%) | 237 (48%) |
| | <i>He is</i> | <i>Internal</i> | 113 | 127 |
| | <i>He is</i> | <i>External</i> | 109 | 138 |
| | Total <i>He is</i> (all orientations) | | 222 (46%) | 265 (54%) |

Table 6-3. Predictors of image choice in Experiment 4: Coefficients from Model 6-2 (significant fixed effects in bold)

| Predictor | Coefficient | Standard Error | <i>p</i> (coefficient = 0) |
|---|-------------|----------------|----------------------------|
| <i>Likelihood of choosing right-handed image</i> | | | |
| (Intercept) | 0.11 | 0.06 | 0.05* |
| Response hand | 0.08 | 0.11 | 0.45 |
| Pronoun | -0.15 | 0.10 | 0.15 |
| Image orientation | 0.13 | 0.12 | 0.26 |
| Image-screen congruency | 1.40 | 0.11 | <.001*** |
| Response hand x pronoun | 0.24 | 0.21 | 0.25 |
| Response hand x image orientation | 0.25 | 0.19 | 0.20 |
| Pronoun x image orientation | 0.21 | 0.19 | 0.29 |
| Response hand x pronoun x image orientation | -0.16 | 0.44 | 0.71 |

Button Press RTs

I checked for trials with button press RTs of under 200 ms (0 trials), and winsorised remaining critical trials so that RTs above or below 2.5 SD for participant's mean response latency were replaced with the upper or lower cutoff value for that participant (0% trials replaced with lower cutoff; 2.25% trials replaced with upper cutoff). Table 6-4 and 6-5 show the mean button press RTs, by condition. There were no significant effects of dominant hand, pronoun, image orientation, or image-screen congruency, and no significant interactions between these variables using Model 6-1 (all $ps > .05$). In addition, participants were not significantly faster to select right-handed images compared with left-handed images ($B = -8.531$, $SE = 14.94$, $p = 0.57$).

Table 6-4. Mean winsorised button press RTs (ms) by image-screen congruency in Experiment 4 (sd in parentheses).

| Response hand | Image-screen congruency | Image choice | |
|---------------|--|-------------------|--------------------|
| | | Left-handed image | Right-handed image |
| Left hand | <i>Right-handed image on right of screen</i> | 1251 (445) | 1228 (413) |
| | <i>Left-handed image on right of screen</i> | 1171 (364) | 1267 (435) |
| Right hand | <i>Right-handed image on right of screen</i> | 1288 (458) | 1174 (405) |
| | <i>Left-handed image on right of screen</i> | 1212 (413) | 1259 (440) |

Table 6-5. Mean winsorised button press RTs (ms) by response hand, pronoun and image orientation in Experiment 4 (sd in parentheses)

| | | Image choice | | | | | | | |
|--------------------------|--|-------------------|------------|------------|------------|--------------------|------------|------------|------------|
| | | Left-handed image | | | | Right-handed image | | | |
| <i>Pronoun</i> | | I am | He is | I am | He is | I am | He is | I am | He is |
| <i>Image orientation</i> | | Internal | Internal | External | External | Internal | Internal | External | External |
| Left hand | | 1223 (414) | 1235 (414) | 1201 (386) | 1135 (358) | 1310 (478) | 1236 (417) | 1212 (395) | 1198 (374) |
| Right hand | | 1260 (444) | 1232 (433) | 1234 (419) | 1219 (424) | 1241 (483) | 1144 (341) | 1190 (442) | 1232 (398) |

6.3.3. Discussion

In Experiment 4, I found that when participants read simple action sentences without knowing which hand will be used to respond, there is no effect of response hand on image choice, and no interaction with pronoun. The effect of image-screen congruency remains significant, suggesting that the paradigm is still sensitive to some effects. Post-hoc simulation-based power analysis were used to check that the design of Experiment 4 had sufficient power to detect an effect of response hand. The effect of response hand is estimated based on the comparison between right-handers in Experiments 2a and 2b ($b_0 = 0.79$, $b_1 = 0.33$). The standard deviation associated with subjects (0.18) and items (0.06) were based on the actual data from Experiment 4. Because these data are analysed using logistic regression, both the coefficients for the fixed effects of response hand, and the standard deviations associated with random effects, are specified on the logit scale. Results of this power analysis showed that Model 6-2 correctly detected the effect of response hand on only 5% of 1000 simulated datasets, indicating an underpowered design. An additional power analysis was therefore conducted, using the following model which collapsed across pronoun and image orientation:

Model 6-3: Response hand + (1 + Response hand| Subject) + (1 + Response hand| Item)

Using this simplified model, the effect of response hand was correctly detected on >99% simulations. The actual data obtained from Experiment 4 were then reanalysed using Model 6-3 to check for any effect of response hand when collapsing across pronoun and image orientation. Results of this additional analysis showed no effect of response hand on the probability of selecting the right-handed image ($b = 0.07$, $SE = 0.09$, $p = .41$), despite the increase in power.

Together, these results support the hypothesis that the effect of response hand in Experiment 2a and 2b are driven by motor resonance present *during* sentence processing. However, the insertion of the coloured screen to cue response hand resulted in a 1500 ms delay between sentence offset and image presentation. If perspective-taking is a short-lived phenomenon, then the 1500 ms gap between sentence and image presentation might nullify any effect. Once the image is presented, the corresponding sentence might no longer be active enough to trigger perspective-taking in the comprehender. In Experiment 5, I tested this possibility by retaining the 1500 ms gap, but having participants know their response hand during sentence processing.

6.4. Experiment 5

In Experiment 5 I checked whether the delay between sentence and image presentation in Experiment 4 compared with Experiment 2a and 2b could explain the different results. I followed exactly the same procedure and importantly, timing, as Experiment 4, with the exception that trials were blocked instead of randomized so that participants knew, while reading the sentence, which hand they would use to respond with. Motor resonance was therefore available to interact with language *during* sentence processing. If the null results in Experiment 4 were the result of a delay between sentence and image presentation, then we should also observe null results in Experiment 5. If the null results in Experiment 4 were the result of motor resonance not being present during sentence presentation, then Experiment 5 should replicate the pattern of results for right-handed participants in Experiment 2a and 2b (i.e., there should be a main effect of response hand and an interaction between response hand and pronoun).

6.4.1. Method

Participants

Sixteen native English speakers (9 females, mean age = 20.9 years, age range = 18 to 34) took part in return for course credit or payment. All participants had normal or corrected to normal vision, and no history of reading difficulties. All participants were right-handed according to the EHI (mean EHI score = 83, EHI score range: 53 to 100).

Materials and design

Materials and design were the same as Experiment 4, with the exception that blue box and green box trials were blocked together rather than randomly presented. This blocking procedure meant that, even though the coloured screen was not presented until after sentence presentation, participants knew their response hand while reading the sentence.

Procedure

The procedure was the same as Experiment 4. No mention of handedness was made prior to testing, and no participant noticed a connection between dominant or response hand and hand in the experimental images.

6.4.2. Results

Analysis

Analyses for error rates were the same as for Experiment 4, using Model 6-1. Analyses for image choice and button press RTs were the same as for Experiment 4, using Model 6-2. In addition, a further analysis with trial as a fixed effect was carried on each block separately, to check for a cumulative effect of trial over the course of a block. For these analyses, trial was added as fixed effect to Model 6-2 to give the following model with simplified random effect structure (the correlation parameter, the random intercepts and

slopes for image-screen congruency, and the by-item random intercepts and slopes for trial and for interactions with item orientation were removed in order to allow the model to converge):

Model 6-4: Pronoun * Image orientation + Trial + Image-screen congruency + (1+ Pronoun * Image orientation + Trial | Subject) + (1+ Pronoun + Image orientation | Item)

Error Rates

Sentence-picture matching

Overall accuracy on selecting the correct image to match the sentence on filler trials was 93%. There were no significant differences in accuracy for first-person sentences (93%) versus third-person sentences (93%); for internally (93%) versus externally (93%) orientated images; for right-handed (93%) versus left-handed (93%) responses; or for image-screen congruent (93%) versus incongruent (93%) trials (all $ps > .05$).

Sentence repetition

Overall accuracy for sentence repetition was 98% (99% on critical trials; 98% on filler trials, ns.). There were no significant differences in accuracy for first-person sentences (98%) versus third-person sentences (99%); for internally (97%) versus externally (99%) orientated images; for image-screen congruent (98%) versus incongruent (99%) trials; or for right-handed (99%) versus left-handed (98%) participants (all $ps > .05$).

Response hand

Accuracy for using the correct response hand was 100%.

Image choice

In total, 1% of critical trials timed-out without a response, and could not be analysed. I checked for trials where participants selected an image within 200 ms of image onset (0 critical trials). Tables 6-6 and 6-7 show the frequency with which participants

selected a right-handed image, by condition. Table 6-8 shows the model coefficients for the likelihood of selecting a right-handed image over the whole experiment, using Model 6-2.

There was a main effect of response hand: participants were significantly more likely to select a right-handed image when using their right hand compared with their left hand. There was no significant effect of pronoun, but the interaction between response hand and pronoun was significant. There was a significant effect of image-screen congruency: in the congruent condition (right-handed image on right side of the screen), participants were more likely to select the right-handed image, and in the incongruent condition (left-handed image on the right side of the screen), participants were more likely to select the left-handed image (i.e., they were more likely to select the image on right side of the screen, regardless of response hand). There was no effect of trial, and no effects of image orientation, as a main effect, or in interaction with other predictors (all $ps > .05$).

Table 6-6. Frequency of right-handed and left-handed image choices by image-screen congruency in Experiment 5 (percentage responses in parentheses)

| Response hand | Image-screen congruency | Image choice | |
|---------------|--|-------------------|--------------------|
| | | Left-handed image | Right-handed image |
| Left hand | <i>Right-handed image on right of screen</i> | 243 (48%) | 263 (52%) |
| | <i>Left-handed image on right of screen</i> | 367 (72%) | 138 (28%) |
| Right hand | <i>Right-handed image on right of screen</i> | 148 (29%) | 360 (61%) |
| | <i>Left-handed image on right of screen</i> | 223 (44%) | 288 (56%) |

Table 6-7. Frequency of right-handed and left-handed image choice by response hand, pronoun, and image orientation in Experiment 5 (percentage responses in parentheses).

| | Pronoun | Image orientation | Left-handed image | Right-handed image |
|------------------------|---------------------------------------|--------------------------|-------------------|--------------------|
| Left-handed responses | I am | Internal | 153 | 100 |
| | I am | External | 165 | 87 |
| | Total I am (all orientations) | | 318 (63%) | 187 (37%) |
| | He is | Internal | 147 | 108 |
| | He is | External | 145 | 106 |
| | Total He is (all orientations) | | 292 (58%) | 214 (42%) |
| Right-handed responses | I am | Internal | 68 | 190 |
| | I am | External | 79 | 171 |
| | Total I am (all orientations) | | 147 (29%) | 361 (71%) |
| | He is | Internal | 116 | 141 |
| | He is | External | 108 | 146 |
| | Total He is (all orientations) | | 224 (44%) | 287 (56%) |

Model 6-4 (incorporating an trial as a fixed effect and by-subjects random effect), was run on each half of Experiment 5, as described above, to check whether participants showed any evidence of cumulative effect of response hand over time (see Table 6-9). For the right hand response block, there was a significant effect of pronoun and image-screen congruency; for the left hand response block, there was a significant effect of image-screen congruency. There was no effect of trial in either block, suggesting that participants did not, for example, become more “left-handed” over the left-handed response block.

Table 6-8. Predictors of image choice in Experiment 5: Coefficients from Model 6-2 (significant fixed effects in bold)

| Predictor | Coefficient | Standard Error | <i>p</i> (coefficient = 0) |
|---|-------------|----------------|----------------------------|
| <i>Likelihood of choosing right-handed image</i> | | | |
| (Intercept) | -0.07 | 0.17 | 0.68 |
| Response hand | 1.21 | 0.39 | .002** |
| Pronoun | 0.19 | 0.19 | 0.30 |
| Image orientation | 0.09 | 0.11 | 0.41 |
| Image-screen congruency | 0.99 | 0.10 | <.001*** |
| Response hand x pronoun | 1.01 | 0.21 | <.001*** |
| Response hand x image orientation | -0.05 | 0.10 | 0.81 |
| Pronoun x image orientation | .033 | 0.20 | 0.09 |
| Response hand x pronoun x image orientation | 0.16 | 0.40 | 0.69 |

Button press RTs

I checked for trials with button press RTs of under 200 ms (0 trials), and winsorised remaining critical trials so that RTs above or below 2.5 SD for participant's mean response latency were replaced with the upper or lower cutoff value for that participant (0% trials replaced with lower cutoff; 1.81% trials replaced with upper cutoff). Tables 6-10 and 6-11 show the mean button press RTs, by condition. I checked for effects of response hand, pronoun, and image orientation on button press RTs as well as trial and image-screen congruency, using Model 6-2. There were no significant effects of pronoun, image orientation, image-screen congruency, and no interactions between these variables (all $ps > .05$). The effect of response hand was marginally significant ($B = -83.67$, $SE = 51.52$, $p = 0.1$) The by-block analysis revealed no significant effects of response hand, pronoun, image orientation, or image-screen congruency (all $ps > .05$), but trial was significant ($p < .001$) for both right-handed and left-handed response blocks.

Table 6-9. Predictors of image choice in Experiment 5 by block: Coefficients from Model 6-4 (significant fixed effects in bold)

| Predictor | RIGHT-HAND RESPONSES | | | LEFT-HAND RESPONSES | | |
|---|----------------------|----------------|----------------------------|---------------------|----------------|--------------------------|
| | Coefficient | Standard Error | <i>p</i> (coefficient = 0) | Coefficient | Standard Error | <i>p</i> (coefficient=0) |
| <i>Likelihood of choosing right-handed image</i> | | | | | | |
| (Intercept) | 1.6 | 0.05 | <.001*** | 1.36 | 0.06 | <.001*** |
| Pronoun | 0.15 | 0.05 | <.01** | -0.06 | 0.05 | 0.25 |
| Image orientation | 0.02 | 0.03 | 0.60 | 0.03 | 0.03 | 0.43 |
| Trial | <.001 | <.001 | 0.40 | <.001 | <.001 | 0.48 |
| Image-screen congruency | 0.16 | 0.03 | <.001*** | 0.24 | 0.03 | <.001*** |
| Pronoun x image orientation | 0.08 | 0.06 | 0.21 | 0.05 | 0.06 | 0.40 |

Table 6-10. Mean winsorised button press RTs (ms) by image-screen congruency in Experiment 5 (sd in parenthesis)

| Response hand | Image-screen congruency | Image choice | |
|---------------|--|-------------------|--------------------|
| | | Left-handed image | Right-handed image |
| Left hand | <i>Right-handed image on right of screen</i> | 1289 (405) | 1095 (436) |
| | <i>Left-handed image on right of screen</i> | 1169 (404) | 1227 (452) |
| Right hand | <i>Right-handed image on right of screen</i> | 1319 (473) | 1255 (448) |
| | <i>Left-handed image on right of screen</i> | 1258 (464) | 1293 (430) |

Table 6-11. Mean winsorised button press RTs (ms) by response hand, pronoun, and image orientation in Experiment 5 (sd in parentheses)

| Image choice | | | | | | | | |
|---------------------------------|------------|------------|------------|------------|--------------------|------------|------------|------------|
| Left-handed image | | | | | Right-handed image | | | |
| <i>Pronoun</i> | I am | He is | I am | He is | I am | He is | I am | He is |
| <i>Image orientation</i> | Internal | Internal | External | External | Internal | Internal | External | External |
| Left hand | 1172 (345) | 1199 (437) | 1248 (457) | 1249 (459) | 1131 (444) | 1177 (435) | 1097 (420) | 1149 (403) |
| Right hand | 1237 (439) | 1298 (438) | 1288 (490) | 1289 (503) | 1244 (404) | 1212 (403) | 1367 (496) | 1257 (436) |

Between experiment comparisons

In order to confirm that the blocking versus random manipulation led to different effects on image choice, I conducted a planned comparison between Experiments 4 and 5. Since image orientation had not been significant in any previous models, it was removed as a fixed effect from the model, and experiment was added as a fixed effect. Analyses were therefore conducted using the following model with simplified random effect structure (the random intercepts and slopes for image-screen congruency were removed in order to allow the model to converge):

Model 6-5: Response hand * Pronoun * Experiment + Image-screen congruency +
(1+Response hand * Pronoun | Subject) + (1+Response hand * Pronoun * Experiment |
Item)

Table 6-12 shows the model coefficients for this comparison. Across both experiments, there was a significant effect of response hand and of image-screen congruency. There was a significant interaction between experiment and response hand: effect of response hand was significantly stronger in Experiment 5 than Experiment 4. The interaction between pronoun and response hand was significant across the two experiments, and the three-way interaction between pronoun, response hand and experiment was also significant.

6.4.3. Discussion

In Experiment 5, right-handed participants were more likely to select right-handed images when responding with their right hand, and left-handed images when responding with their left hand. This preference for images matching hand-in-use was stronger following first-person compared with third-person sentences. These results contrasted with the null results in Experiment 4. A planned between-experiment comparison revealed that the three way

interaction between response hand, pronoun, and experiment was also significant, suggesting that the null results in Experiment 4 cannot be attributed to increased time lapse between sentence offset and image presentation, compared with Experiments 2a and 2b.

Table 6-12. Predictors of image choice across Experiments 4 and 5: Coefficients from Model 6-5 (significant fixed effects in bold)

| Predictor | Coefficient | Standard Error | <i>p</i> (coefficient = 0) |
|---|--------------|----------------|----------------------------|
| <i>Likelihood of choosing right-handed image</i> | | | |
| (Intercept) | 0.09 | 0.04 | 0.03* |
| Response hand | -0.62 | 0.22 | .004** |
| Pronoun | 0.03 | 0.10 | 0.74 |
| Experiment | 0.03 | 0.08 | 0.70 |
| Image-screen congruency | -1.2 | 0.07 | <.001*** |
| Response hand x Pronoun | -0.55 | 0.22 | 0.01* |
| Response hand x Experiment | 1.07 | 0.44 | 0.001** |
| Pronoun x Experiment | -0.35 | 0.20 | 0.07 |
| Response hand x Pronoun x Experiment | 1.43 | 0.43 | <.001*** |

In summary, Experiment 5 provides a within-subjects replication of Experiments 2a and 2b, with significant effects of response hand, response hand by pronoun interaction, and image-screen congruency. Experiment 4 on the other hand, showed only an effect of image-screen congruency. Experiments 4 and 5 differed in that in Experiment 4, we cued left or right response hand randomly after sentence presentation; but in Experiment 5, we blocked left and right response hands so that participants knew their response hand during sentence processing. The significant effect of this manipulation suggests that the effect of hand-in-use and its interaction with pronoun occur only when participants know, during sentence

processing, which hand they will use to respond with. In other words, response hand and pronoun affect action sentence interpretation only when participants are able to plan a motor response with their response hand, during sentence processing.

However, an alternative explanation is that it is not planned response hand, but previous response hand, that is responsible for the effect of response hand on sentence interpretation. Once again, there was again no effect of trial on the likelihood of selecting a right-handed image in Experiment 5, indicating the participants do not become more left handed over the course of a left-handed block. However, blocking response hand meant that, as in Experiments 2a and 2b, response hand on the previous trial coincided with planned response hand on the present trial. Therefore, the results in Experiments 2a, 2b and 5 may have been caused by response hand on the previous trial, rather than participants planning to respond with a given hand during sentence comprehension. Experiment 6 was therefore designed to distinguish between the effects of response hand on the previous trial (retrospective simulation) and planned response hand on the current trial (prospective simulation).

6.5. Experiment 6

In Experiment 6 I checked whether the effect of current motor context in Experiments 2a, 2b and 5 could be attributed to previous response hand (retrospective simulation) or to planned response hand (prospective simulation). We did this by randomly cueing response hand prior to sentence presentation. Presenting the cues before sentence presentation meant that, like Experiment 5, participants knew their response hand during sentence processing. The randomization process meant that, unlike Experiment 5, previous response hand could be separated from planned response hand. Other than randomizing the response hand cues, procedure was identical to Experiment 5. If the results in Experiments 2a, 2b and 5 were due to previous response hand, then there should be an effect of previous response hand, and no effect of current (planned) response

hand. If the results in Experiments 2a, 2b and 5 were due to planned response hand, then there should be an effect of current (planned) response hand, and no effect of previous hand.

6.5.1. Method

Participants

Sixteen native English speakers (8 females, mean age = 19.2 years, age range = 17 to 22) took part in return for course credit or payment. All participants had normal or corrected to normal vision, and no history of reading difficulties. All participants were right-handed according to the EHI (mean EHI score = 87, EHI score range: 58 to 100).

Materials and design

Materials and design were the same as Experiment 4.

Procedure

The procedure was the same as Experiment 4, with the exception that the randomly selected blue or green screen cues appeared before the sentence, rather than after it. This manipulation meant that left and right hand responses were randomized, but participants knew which hand they would respond with while reading the sentence. No mention of handedness was made prior to testing, and no participant noticed a connection between dominant or response hand and hand in the experimental images.

6.5.2. Results

Analysis

Analyses were the same as for Experiment 4, with the exception that for image choice and button press analyses, we included response hand from the previous trial as a fixed effect, giving the following model with simplified random effect structure (the

correlation parameter, the random intercepts and slopes for image-screen congruency and previous response hand, and the by-item random intercepts and slopes for interactions with item orientation were removed in order to allow the model to converge):

Model 6-6: Response hand * Pronoun * Image orientation + Previous response hand + Image-screen congruency + (1 | Subject) + (1 | Item) + (0+Response hand * Pronoun * Perspective | Subject) + (0+Response hand * Pronoun + Perspective | Item)

Error Rates

Sentence-picture matching

Overall accuracy on selecting the correct image to match the sentence on filler trials was 95%. There were no significant differences in accuracy for first-person sentences (95%) versus third-person sentences (95%); for internally (96%) versus externally (95%) orientated images; for right-handed (95%) versus left-handed (96%) responses; or for image-screen congruent (96%) versus incongruent (95%) trials (all $ps > .05$).

Sentence repetition

Overall accuracy for sentence repetition was 98% (99% on critical trials; 98% on filler trials, ns.). There were no significant differences in accuracy for first-person sentences (98%) versus third-person sentences (99%); for internally (97%) versus externally (99%) orientated images; for image-screen congruent (98%) versus incongruent (99%) trials; or for right-handed (99%) versus left-handed (98%) participants (all $ps > .05$).

Response hand

Overall accuracy on using the correct hand to respond on critical trials was 96% (96% on critical trials; 96% on filler trials, ns.). There were no significant differences in accuracy for first-person sentences (96%) versus third-person sentences (96%); for internally (96%) versus externally (96%) orientated images; for right-handed (97%) versus left-handed

(96%) responses; or for image-screen congruent (96%) versus incongruent (96%) trials (all $ps > .05$).

Image choice

In total, 1% of critical trials timed-out without a response, and could not be analysed. I checked for trials where participants selected an image within 200 ms of image onset (0 critical trials). Tables 6-13 and 6-14 show the frequency with which participants selected a right-handed image, by condition. Table 6-15 shows the model coefficients for the likelihood of selecting a right-handed image, using Model 6-6. There was a main effect of response hand: participants were significantly more likely to select a right-handed image when using their right hand compared with their left hand. There was no significant effect of pronoun, but the interaction between response hand and pronoun was significant. There was a significant effect of image-screen congruency, with participants being more likely to select the congruent image than the incongruent image. There was no effect of the hand used to respond on a previous trial, and no effects of image orientation, as a main effect, or in interaction with other predictors (all $ps > .05$).

Table 6-13. Frequency of right-handed and left-handed image choices by image-screen congruency in Experiment 6 (percentage responses in parentheses)

| Response hand | Image-screen congruency | Image choice | |
|---------------|--|-------------------|--------------------|
| | | Left-handed image | Right-handed image |
| Left hand | <i>Right-handed image on right of screen</i> | 320 (61%) | 203 (39%) |
| | <i>Left-handed image on right of screen</i> | 310 (60%) | 209 (40%) |
| Right hand | <i>Right-handed image on right of screen</i> | 156 (30%) | 370 (70%) |
| | <i>Left-handed image on right of screen</i> | 228 (44%) | 291 (56%) |

Table 6-14. Frequency of right-handed and left-handed image responses by response hand, pronoun, and image orientation in Experiment 6 (percentage responses in parentheses).

| | Pronoun | Image orientation | Left-handed image | Right-handed image |
|------------------------|---------------------------------------|-------------------|-------------------|--------------------|
| Left-handed responses | I am | Internal | 170 | 88 |
| | I am | External | 155 | 105 |
| | Total I am (all orientations) | | 325 (63%) | 193 (37%) |
| | He is | Internal | 161 | 105 |
| | He is | External | 144 | 114 |
| | Total He is (all orientations) | | 305 (58%) | 219 (42%) |
| Right-handed responses | I am | Internal | 84 | 177 |
| | I am | External | 82 | 180 |
| | Total I am (all orientations) | | 166 (32%) | 357 (68%) |
| | He is | Internal | 109 | 155 |
| | He is | External | 109 | 149 |
| | Total He is (all orientations) | | 218 (42%) | 304 (58%) |

Table 6-15. Predictors of image choice in Experiment 6: Coefficients from Model 6-6 (significant fixed effects in bold).

| Predictor | Coefficient | Standard Error | p (coefficient = 0) |
|--|--------------|----------------|---------------------|
| Likelihood of choosing right-handed image | | | |
| (Intercept) | -0.03 | 0.21 | 0.89 |
| Response hand | -1.26 | 0.30 | <.001*** |
| Pronoun | -0.01 | 0.18 | 0.94 |
| Image orientation | -0.20 | 0.12 | 0.10 |
| Previous response hand | -0.09 | 0.10 | 0.37 |
| Image-screen congruency | 0.37 | 0.11 | <.001*** |
| Response hand x pronoun | -1.01 | 0.39 | 0.007** |
| Response hand x image orientation | -0.39 | 0.20 | 0.06 |
| Pronoun x image orientation | -0.19 | 0.20 | 0.35 |
| Response hand x pronoun x image orientation | -0.20 | 0.44 | 0.64 |

Prompted by the null effect of pronoun for left-handed responses in Experiment 5, I conducted an additional analysis on Experiment 6, to check for an effect of pronoun in right-handed and left-handed responses separately using the following model with simplified random effect structure (the correlation parameter, the random intercepts and slopes for image-screen congruency and previous response hand, and the by-item random intercepts and slopes for interactions with item orientation were removed in order to allow the model to converge):

Model 6-7: Pronoun * Image orientation + Previous response hand + Image-screen congruency + (1 | Subject) + (1 | Item) + (0+ Pronoun * Perspective | Subject) + (0+ Pronoun + Perspective | Item)

Table 6-16 shows the results of these analyses. For right hand responses, there was a significant effect of pronoun and of image-screen congruency; for left hand responses block, there were no significant effects on image choice.

Button press RTs

I checked for trials with button press RTs of under 200 ms (0 trials), and winsorised remaining critical trials so that RTs above or below 2.5 SD for participant's mean response latency were replaced with the upper or lower cutoff value for that participant (0% trials replaced with lower cutoff; 1.34% trials replaced with upper cutoff). Tables 6-17 and 6-18 show the mean button press RTs, by condition. I checked for effects of response hand, pronoun, and image orientation on button press RTs as well as previous response hand and image-screen congruency, using Model 6-5. There were no significant effects of previous response hand, pronoun, image orientation, previous response hand or image-screen congruency, and no significant interactions between these variables (all $ps > .05$).

Table 6-16. Predictors of image choice in Experiment 6 by response hand: Coefficients from Model 6-7 (significant fixed effects in bold)

| Predictor | RIGHT-HAND RESPONSES | | | LEFT-HAND RESPONSES | | |
|---|----------------------|----------------|----------------------------|---------------------|----------------|--------------------------|
| | Coefficient | Standard Error | <i>p</i> (coefficient = 0) | Coefficient | Standard Error | <i>p</i> (coefficient=0) |
| <i>Likelihood of choosing right-handed image</i> | | | | | | |
| (Intercept) | 1.63 | 0.04 | <.001*** | 1.40 | 0.06 | <.001*** |
| Pronoun | 0.10 | 0.03 | 0.002** | -0.05 | 0.05 | 0.36 |
| Image orientation | -.001 | 0.04 | 0.99 | -0.06 | 0.03 | 0.06 |
| Previous response hand | -0.02 | 0.03 | 0.32 | -0.001 | 0.03 | 0.09 |
| Image-screen congruency | 0.16 | 0.03 | <.001*** | -0.01 | 0.03 | 0.59 |
| Pronoun x image orientation | -0.01 | 0.07 | 0.85 | -0.01 | 0.06 | 0.80 |

Table 6-17. Mean winsorised button press RTs (ms) by image-screen congruency in Experiment 6 (sd in parenthesis)

| Response hand | Image-screen congruency | Image choice | |
|---------------|--|-------------------|--------------------|
| | | Left-handed image | Right-handed image |
| Left hand | <i>Right-handed image on right of screen</i> | 1349 (452) | 1287 (468) |
| | <i>Left-handed image on right of screen</i> | 1392 (452) | 1298 (470) |
| Right hand | <i>Right-handed image on right of screen</i> | 1421 (450) | 1257 (447) |
| | <i>Left-handed image on right of screen</i> | 1369 (496) | 1361 (414) |

6.5.3. Discussion

In Experiment 6, right-handed participants were cued to respond with their left or right hand before reading a sentence describing a manual action, and selecting a right-handed or left-handed image to match the sentence. Participants were more likely to select a right-handed image when they responded using their right hand, and more likely to select a left-handed image when they responded with their left hand. This effect of hand-in-use interacted with pronoun: participants were more likely to select an image that was congruent with their response hand following first-person than third-person sentences. Experiment 6 therefore provided a further within-subjects replication of the hand by pronoun interaction in Experiments 2a and 2b. In addition, participants were not more likely to select a right-handed image having used their right-hand on a previous trial. Only the hand that participants had been planning to use when reading the sentence had an effect on the likelihood of choosing a right-handed image. This pattern of results, coupled with the lack of a cumulative effect of trial in Experiment 5, implies that the motor resonance implicated in action language processing comes from motor planning rather than immediate short-term motor experience.

Table 6-18. Mean winsorised button press RTs (ms) by response hand, pronoun, and image orientation in Experiment 6 (sd in parentheses)

| Image choice | | | | | | | | |
|--------------------------|------------|------------|------------|------------|--------------------|------------|------------|------------|
| Left-handed image | | | | | Right-handed image | | | |
| Pronoun | I am | He is | I am | He is | I am | He is | I am | He is |
| Image orientation | Internal | Internal | External | External | Internal | Internal | External | External |
| Left hand | 1294 (408) | 1334 (479) | 1420 (431) | 1446 (479) | 1182 (434) | 1422 (526) | 1262 (442) | 1286 (439) |
| Right hand | 1387 (463) | 1383 (455) | 1413 (531) | 1382 (475) | 1310 (467) | 1328 (445) | 1283 (410) | 1293 (420) |

6.6. General Discussion

In three experiments, I manipulated at what point in time (before or after sentence processing) comprehenders were able to plan a motor response on a given trial. In Experiment 4, participants were shown a randomly selected response hand cue after sentence processing. There was no effect of response hand on image choice, and no interaction between response hand and pronoun. In Experiment 5, participants were shown a blocked response hand cue after sentence processing. There was a significant effect of response hand on image choice, and this effect was stronger following first-person sentences (however, separate analyses of the right-handed and left-handed halves of the experiment found a significant effect of pronoun only in right-handed trials). There was no effect of trial on image choice, suggesting that the results are not due to cumulative short-term motor experience caused by the blocking procedure. The between experiments comparison of Experiments 4 and 5 was also significant. In Experiment 6, participants were shown a randomly selected response hand cue before sentence processing. There was a significant effect of response hand on image choice, and this effect was stronger following first-person sentences. There was no effect of the hand used to respond on the previous trial.

Experiments 5 and 6 provided within-subjects replications of Experiments 2a and 2b, and therefore support the conclusions from Chapter 5 that language comprehenders tend to adopt an embodied agent's perspective on action language, and that this tendency is more pronounced for language that is potentially self-referential than for language that unambiguously refers to a third party. However, further analyses of both experiments show a significant effect of pronoun only in the right-handed trials (contra. Experiment 2b), suggesting that the effects of perspective taking may be stronger when dominant hand and response hand coincide, than when they do not. The contrasting results of Experiments 4 and 5 extend the findings from Chapter 5, by suggesting that hand-in-use

only affects image choice when participants are aware of their response hand during sentence processing; no effects are observed when response hand is specified after sentence processing. In other words, motor resonance must be present during sentence processing. The null effects in Experiment 4 cannot be attributed to delay between sentence presentation and image presentation (cf. Experiment 5), or the random cueing of response hand (cf. Experiment 6). Moreover, the results from Experiment 6 (a significant effect of response hand, and no effect of the hand used to respond on the previous trial) suggest that the source of this motor resonance is the motor planning of upcoming actions, rather than residual activation from short-term motor experience.

6.6.1. *Prospective versus retrospective simulation*

In Chapter 2 (section 2.4), I discussed evidence that embodied representations in language are extremely flexible, and that this flexibility may account for some of the reported inconsistencies in the literature. For example, the degree or directions of embodied representations appears to be affected by subject pronoun (Brunyé et al., 2011; Papeo et al., 2011), sentence tense (Bergen & Wheeler, 2010), timing of stimulus materials (Borreggine & Kaschak, 2006; Diefenbach et al., 2013), and strategy employed on previous tasks (Papeo et al., 2012).

The influence of previous tasks on language processing is of particular interest to this thesis, in which we aimed to distinguish retrospective simulation (simulating past actions) from prospective simulation (planning future events; see section 2.4.2). In Chapter 5, I presented evidence that people interpret manual action sentences as though they were performing the action themselves, with the caveat that this embodied agent perspective appears to be grounded in hand-in-use rather than dominant hand. In other words, the current motor context overrides long-term motor experience. In the present chapter, I aimed to define the concept of current motor context more exactly: is interpretation of manual action sentences affected by the hand used to respond on a

previous trial (compatible with retrospective simulation), or by the hand the comprehender is planning to use (compatible with prospective simulation)? My results suggest that it is planned actions, rather than recently performed actions, that underpin the motor resonance implicated in action language comprehension. In this sense, the results are in line with recent suggestions that embodied action representations play a predictive role, rather than purely simulating past events (Borghi, 2013; Rueschemeyer & Bekkering, 2013).

The suggestion that motor resonance in action observation has a predictive function, might explain why people are better at recognizing their own actions than those of other people (Repp & Keller, 2010; Repp & Knoblich, 2004), namely, that the recognition advantage for self-generated actions is a result of our improved ability to predict our own actions (see also Knoblich & Flach, 2001; Knoblich & Flach, 2003). A predictive role for simulation might also explain why experts show increased evidence of increased motor resonance compared to novices during action observation (e.g., Calvo-Merino, Glaser, Grèzes, Passingham, & Haggard, 2005; Calvo-Merino, Ehrenberg, Leung, & Haggard, 2010); the experts may be better at *predicting* actions (see Aglioti, Cesari, Romani, & Urgesi, 2008). Note that in the case of experts, the predictive advantage is founded on those individual's long-term motor experience, underlying the fact that prospective simulation builds upon retrospective simulation to predict future actions (see also discussion of motor prototypes in section 2.5.4).

Motor prototypes are central to routine actions – that is, actions or sequences of actions that can be executed without a high degree of concentration, such as getting out of bed or driving to work (Botnivik & Plaut, 2001). The critical sentences in Experiments 1 – 6 describe routine manual actions such as buttering bread or slicing a tomato (although it is true that some actions used as stimuli, such as sewing a shirt, may not be routine to all language comprehenders). Throughout the thesis, and in line with the embodiment literature, I have considered motor resonance to involve activation of those areas of the

motor cortex involved in performing a particular action. However, the motor cortex is not the only area of interest in motor control. Evidence suggests that the cerebellum plays an important role in learning routine actions (Dayan & Cohen, 2011; Thach, 1998). The extent to which motor representations of routine actions, once learned, are stored in the cerebellum versus the motor cortex remains controversial (see, for example, Ashby, Turner, & Horvitz, 2010; Doyon & Benali, 2005; Houk & Wise, 1995). Hikosaka et al. (1999) proposed that the individual sub-movements entailed by an action are controlled by the relevant body-specific mechanisms, but that the cerebellum is responsible for successfully coordinating these sub-routines into an action. Resolving this debate is beyond the scope of this thesis, and one advantage of behavioural paradigms is that they do not depend on prior assumptions about the neural localization of motor representations. However, the extent to which routine actions involve the motor cortex versus the cerebellum is clearly relevant to researchers aiming to provide neuroimaging evidence of embodied representation in language

Interestingly, evidence for a link between retrospective and prospective simulation is provided by the fact that, as well as its involvement in learnt routine actions, the cerebellum has also been heavily implicated in sensorimotor prediction. For example, Wolpert, Miall, and Kawato (1998) suggest that forward models of action prediction (see section 2.4.2) are instantiated in the cerebellum (see also Ito, 2008). The cerebellum's predictive role also appears to extend to language comprehension. In a recent paper, Lesage, Morgan, Olson, Meyer, & Miall (2012) found that applying TMS to the cerebellum prevented language comprehenders from being able to produce predictive eye movements in a visual world paradigm. Participants listened to sentences such as *the man will sail the boat* while viewing a display with one target object (predictable from the verb in the sentence, e.g., *boat*), and three distractor objects (not predictable from the verb in the sentence, e.g., *bird, mountain, car*). Participants in this paradigm typically show increased fixations to the target object before it is mentioned in the sentence (Altmann &

Kamide, 1999). However, participants undergoing TMS in Lesage et al.'s study showed significant delay in eye fixations, compared with control groups, suggesting that they were unable to predict upcoming words. Future work should therefore investigate to what extent the mechanisms of retrospective and predictive simulation rely on shared resources, with particular focus on the role of the cerebellum.

6.6.2. *Is dominant hand defined by hand-in-use?*

Results from this chapter suggest that, when participants are planning an action during sentence processing, this action plan interacts with sentence processing to affect their interpretation of the sentence. At first glance, these results appear to be in conflict with research demonstrating an effect on language processing of long-term motor experience (e.g., Willems, Hagoort, & Casasanto, 2010) and short-term motor experience (e.g., Glenberg et al., 2008). In fact, what these results suggest is that when retrospective simulation (e.g., dominant hand) is placed in direct competition with prospective simulation (e.g., hand-in-use), the effect of prospective simulation may cancel out any effect of retrospective simulation.

The situation in which prospective and retrospective simulation conflict is presumably quite rare. For example, in daily life, the hand typically used to perform an action in the past will almost always coincide with the hand with which you plan to perform future actions. If previous motor experience did not provide a general heuristic for the way in which future actions would be performed, then canonical affordances (Borghi & Riggio, 2009) would have no use. In fact, canonical affordances are useful exactly *because* they provide a guide for the way in which actions are likely to occur at a later date, based on how they have occurred previously. Therefore, one possible conclusion from Experiments 1- 6 is that the context of current task to some extent determines, at a behavioural level at least, hand dominance. Your dominant hand in a given situation is the hand you are producing motor plans for, in that situation. In the

majority of cases, the hand you are planning to use will be determined by the hand you have previously used over your life, or the hand you have been trained to use before the task; but when current planning is placed in opposition to previous experience, current planning wins out.

However, there is some reason to suspect that the effect of dominant hand may not be wholly overridden, but rather mitigated, by the effect of hand-in-use. Notably, further analyses in Experiments 5 and 6 revealed an effect of pronoun in right-handed responses, but not in left-handed responses. This finding raises the possibility that, as well as a main effect of hand-in-use, our participants' dominant hand might influence the sensitivity of action simulations to differences in self-referential and non-self-referential language. Perhaps, for example, when current motor plan and long-term experience coincide, the resulting simulation is more strongly coded for agency, compared to when current motor plan and long-term experience of a particular action are in conflict.

6.6.3. *Image-screen congruency*

In five experiments (Experiment 2a, 2b, 4, 5, and 6) I found that people were more likely to select an image when the hand in the image was congruent with the spatial lay out of the images. For example, right-handed images were more likely to be selected when they were presented on the right hand side of the screen, and left-handed images were more likely to be selected when they were presented on the left-hand side of the screen. These findings accord with an extensive body of work on spatial compatibility tasks (e.g., the Simon effect; see Hommel, 2001 for a review). Rubichi and Nicoletti (2006) found evidence to suggest that the Simon effect is moderated by an attentional bias: right- and left-handers performing a Simon task showed a stronger effect in the right visual hemifield and left visual hemifield, respectively. However, when participants performed the same task with their hands crossed, the effects reversed: right-handers showed a stronger effect in their left visual field, and left-

handlers showed a stronger effect in their right visual field. Rubichi and Nicoletti explain their findings by positing an attentional bias towards the dominant hand.

Numeric count of responses in Experiments 4-6 suggest that the effect of image-screen congruency altered depending on participant's response hand: participants using their left hand tended to select a left-handed image more often when this image appeared on the left of the screen than the right of the screen, and participants using their right hand tended to select a right-handed image more often when this image appeared on the right of the screen than the left of the screen. These effects were not tested for significance, but they suggest that image-screen congruency may have interacted with response hand; perhaps the use of a single response hand results in an attentional bias towards that particular hand. More generally, the possible interaction between spatial coding and action simulation is a potentially fruitful avenue of future research.

6.6.4. Strategic processing

Recall from section 2.4.1 that Papeo et al., (2012) found that strategy on a previous task (motor versus visual mental rotation strategy) affected the degree of activation in a subsequent word reading task; they found increased activation when participants had employed a motor strategy on the mental rotation task. These results are in line with work suggesting that language comprehenders show increased levels of motor activation when processing words such as *tennis ball*, when they are asked to think about how a tennis ball is used, compared with its physical appearance (van Dam, van Dijk, Bekkering, & Rueschemeyer, 2012; see also van Dam, Rueschemeyer, Lindemann, & Bekkering, 2010 for compatible findings in a behavioural paradigm). One possible explanation for the results in Experiments 2a - 6 therefore, is that participants were strategically (and consciously) selecting the image that matched their response hand.

There are several reasons to be wary of such an explanation. First, over five experiments, only two participants reported noticing any connection between the handedness

of the images and the hand with which they responded, and these participants were replaced. There was no indication from debriefing that the remaining participants thought the experiment was anything other than a working memory task, in which the real task was correct repetition of the sentences. Second, such a consciously strategic explanation cannot explain the interaction with subject pronoun across 4 different experiments, or the sudden absence of this strategy in Experiment 4 (despite the fact that the effect of image-screen congruency remained constant). An explanation based on conscious, strategic choice based solely on image handedness would not predict any difference between first- and third-person sentences, and it would predict a similar strategy to be in place for Experiment 4, contrary to our findings.

The data from the thesis so far, in particular the fact that interpretation of action language sentences seems so context dependent (in terms of both linguistic features, such as subject pronoun, and motor features, such as current action planning) seems to count against the idea that embodied representations are automatically involved in semantic meaning (Pulvermüller, 1999, 2005). As Rueschemeyer and Bekkering (2013) point out, the strength of language (not to mention our action system) is in its flexibility. It would be intuitively implausible to think that embodied representations in action language were *automatic* in the sense of being unaffected by context; however, they may still be automatic in Heyes' (2011) sense of being beyond the conscious intentions of the comprehender.

6.7. Summary

In this chapter, I found that the embodied agent's perspective on action language only occurs when motor resonance is present during sentence processing, thus counting against explanations which view motor activation as a downstream by-product of comprehension. I replicated the interaction between response hand and pronoun from

Chapter 5, implying that the effect of motor resonance on sentence comprehension is stronger for self-referential language. Furthermore, I found that the source of this motor activation appears to be planned actions, rather than previously performed actions. This finding is in line with more recent views on motor resonance as having a predictive role.

7. A FUNCTIONAL ROLE FOR ACTION PLANNING?

7.1. Overview of the chapter

In Chapter 5, I reported evidence from three experiments suggesting that people adopt an embodied agent's perspective on action language sentences, and that this perspective is grounded in hand-in-use (current motor context) rather than dominant hand (long-term motor experience). In Chapter 6, I reported evidence from three experiments showing that the key feature of current motor context is the hand participants are planning to use (action planning), rather than the hand used to respond on a previous trial (short-term motor experience). In both chapters, I reported experiments in which the effect of hand-in-use is stronger following first-person sentences than third-person sentences, indicating an interaction between action planning and language processing. I also found that the effect of hand-in-use only occurs when participants know their hand-in-use during sentence processing: delaying knowledge of response hand until sentence processing is complete resulted in no effect of response hand or of pronoun. The time course suggested by these results (i.e., that motor resonance must be present during sentence processing for effects to occur) is compatible with a functional role for action planning in language comprehension.

However, in the experiments described above, sentence processing was measured using a manual action response. Therefore, the above results, although compatible with an effect of action on language, might also indicate an effect of language on action. In this chapter, I report three experiments investigating whether there is a causal effect of action-planning based motor resonance on online language processing. I do so by testing whether planning to perform a particular action (e.g., pressing a button pad with a palm or with a single index finger) affects the time taken to make spoken sensibility judgements on sentences describing congruent or incongruent actions. In Experiments 7 and 8, I find no effect of planned action or of response hand on sensibility judgement accuracy or latencies.

In Experiment 9, when participants are expressly instructed to decide whether their planned action is congruent or incongruent with the action described in the sentence, I find an effect of response hand on accuracy, but not reaction time.

7.2. Introduction

Results from Chapter 6 suggested that, in a forced-choice sentence-picture matching task, motor resonance needs to be present during sentence processing in order for there to be effects of hand-in-use and of pronoun. This evidence appears to undermine suggestions that embodied effects in language might be a downstream by-product of comprehension, perhaps based on post-comprehension motor imagery (Hickok, 2010; Mahon & Caramazza, 2008). Such *downstream* explanations would not predict that motor resonance must be present during comprehension for effects to occur. However, as we noted in Chapter 6, evidence for a specified temporal order, in which motor resonance must precede, or at least be concurrent with, sentence comprehension, does not constitute evidence for a causal relationship between the action and language systems. For example, both processes might begin at the same time and then continue without any further interaction (e.g., Mahon & Caramazza, 2008). The interaction between hand-in-use and pronoun in Chapters 5 and 6 suggests that motor resonance and language comprehension do indeed interact with one another, but this still does not necessitate a causal role of action on language.

7.2.1. Measuring language processing

One reason that evidence from Chapters 5 and 6 is insufficient to show a causal relationship is that Chapters 5 and 6 – like the majority of research in this field – investigated the relationship between action and language by measuring differences in motor responses among conditions involving linguistic manipulations. In other words, the dependent variable was a motor response, and the independent variable was language. The use of motor response as an outcome variable is pervasive throughout the embodiment

literature. For example, behavioural researchers have measured button release RTs (Borreggine & Kaschak, 2006; Glenberg & Kaschak, 2002; Kaschak & Borreggine, 2008), button press RTs (Aravena et al., 2010; Borghi et al., 2004), rotation times in self-paced reading (Taylor et al., 2008; Zwaan & Taylor, 2006), and differences in kinematic reaching movements (Boulenger et al., 2006; Nazir et al., 2008). With the exception of kinematic tasks, the motor responses are taken as a proxy for measure of language processing time. In typical psycholinguistic tasks, where the effect of the motor response is assumed to be constant across conditions, the use of a motor proxy to measure linguistic processing is not problematic. However, when the experimental conditions are designed to manipulate the relationship between motor activation and language processing, the use of motor responses as a measure of sentence processing time poses a specific problem: we cannot assume, as we would in wider psycholinguistics, that differences in reaction time reflect straightforward differences in linguistic processing. Rather, they might also reflect differences in action planning or action execution caused by interaction between the language and motor systems.

In order to satisfactorily demonstrate an effect of motor resonance on language processing therefore, we must use action as the independent variable, and language as the dependent variable. In the present chapter, I will explicitly test for an effect of motor resonance on sentence comprehension, by measuring voice onset times following manipulating of planned manual actions. Although voice onset, like any overt movement, requires the motor system to plan and execution actions, it does not require the execution or planning of manual actions. Therefore, by investigating voice onset times to sentences describing manual actions, we can avoid any confound between the effector used in response planning and execution, and the effector implicated in the linguistic stimuli.

7.2.2. Evidence for an effect of action on language

As noted above, with a few notable exceptions, the majority of the embodiment literature demonstrates an effect of language on action, rather than of action on language. In

Chapter 2, I outlined some of those exceptions, which appear to show an effect of action on language, rather than the other way around. For example, a TMS study by Willems and colleagues found that, following TMS stimulation of the hand motor area, participants were faster to respond to manual action verbs in a lexical decision task than to non-manual action verbs (Willems et al., 2011). Similarly, Glenberg, Sato and Cattaneo (2008) found that a ~20 minute training period of moving towards or away from the body affected response times on sensibility judgements on sentences describing movements towards or away from the body in an adapted ACE-paradigm. However, the fact that both studies only measured responses of right-handed participants means that we do not know whether these studies showed comprehenders adopting an embodied agent's perspective, versus that of an embodied observer.

Moreover, these studies cannot tell us anything about the amount of detail present in motor representations of action language. For example, the results from Experiments 1 – 6 imply that participants adopt an embodied agent's perspective, in so much as they interpret manual action sentences as though they were performing a manual action with the hand they are currently planning to use, but it is unclear whether comprehenders simulate performing the *specific* manual action described in the sentence, or merely that they simulate performing a manual action. In this chapter, I will investigate the effect of congruency between the specific handshape (POKE, with only index finger extended; PALM, with entire hand flat and all fingers extended) implied by a manual action sentence, and the handshape required by an action that participants are planning to perform. In this way, the studies in this chapter will elucidate (a) whether there is evidence of a functional role of motor resonance on online language processing; and (b) whether the perspective adopted by comprehenders specifies what action is performed.

In a study published before the recent upsurge in interest in embodied approaches to language, Klatzky, Pellegrino, McCloskey and Doherty (1989) primed participants by showing them one of four sets of word pairs, where each word pair depicted a different

handshape: *touch – finger* [POKE handshape], *grasp – finger* [PINCH handshape], *touch – hand* [PALM handshape]; and *grasp – palm* [CLENCH handshape]. Following presentation of the prime, participants judged the sensibility of verb-noun phrases such as *Aim a dart* (sensible) and *Aim a band-aid* (nonsense). Results showed that participants were faster to make the sensibility judgements when the handshape required to perform the action in the sentence was congruent with the handshape depicted in the prime. For example, participants were faster to judge *Aim a dart* as sensible following the prime *grasp – finger*, which depicted the pinch handshape, compared with other primes.

Importantly, in Klatzky et al.'s study, the independent variable (the prime) was linguistic, while the dependent variable was motoric: participants responded by pressing the *p* or *q* keys on a keyboard (i.e., by making an overt POKE action). The results therefore could therefore be interpreted as demonstrating an effect of language on action, rather than of action on language. However, Klatzky et al.'s results do suggest that an effect of motor activation on language processing might be fine-grained enough to code for the specific handshape required to perform a described action. In the present chapter, I test this possibility by reworking the study by Klatzky et al. so that the independent variable (planned action) is motoric, and the dependent variable (spoken sensibility judgement) is linguistic (to the extent that an overt movement is required to make to vocal response, this response is none the less independent of that part of the motor system implicated in manual actions). By removing manual action from the response criteria, I will search for evidence of a functional role of motor resonance on language processing.

7.2.3. Long-term motor experience, revisited

The results from Experiments 1 - 6 suggest that the effect of long-term motor experience (i.e., dominant hand) is overridden by the current motor context (specifically, the hand a comprehender is planning to use). In Experiment 3, when no overt manual response was required in the task, there was no evidence of an effect of dominant hand. However, in

this chapter I will continue to test for an effect of dominant hand. There are several reasons for doing so. First, although Experiments 5 and 6 showed overall interactions between respond hand and pronoun, separate analyses of right-handed and left-handed responses showed an effect of pronoun only on right-handed responses (i.e., those responses that used the participants' dominant hand). This finding suggests that long-term motor experience may, in fact, interact with simulations based on current action plan, such that the case in which dominant hand and hand-in-use coincide results in more detailed and sensitive action simulations than the case in which dominant hand and hand-in-use conflict.

Second, as noted above, the forced-choice sentence-picture matching paradigm used in Chapters 5 and 6 required an overt manual response. The sensibility judgement paradigm in the present chapter requires no manual response to linguistic stimuli; however, unlike Experiment 3, participants will be planning a manual response during sentence processing. Therefore, the hand motor area will be active during sentence processing (unlike Experiment 3), but sentence processing will not be measured using a manual response (unlike Experiments 4-6). In this "intermediate" state, it is possible that an underlying effect of long-term motor experience on language processing, as suggested by fMRI research (Willems, Hagoort, et al., 2010), might become apparent.

7.2.4. Sentence sensibility judgements

Sensibility judgement tasks are used throughout psycholinguistics, and are assumed to require a reasonably deep level of semantic processing from the comprehender (Louwerse & Jeuniaux, 2008). Sensibility judgements have been particularly widely used to investigate Action-Sentence Compatibility Effects (Borreggine & Kaschak, 2006; Gianelli et al., 2011; Glenberg & Kaschak, 2002; Glenberg, Sato, & Cattaneo, 2008; Kaschak & Borreggine, 2008). Such studies typically do not report the results for non-sensible sentences, although Klatzky et al. (1989) report that the priming effect in their study was limited to sensible sentences. According to Glenberg (2007), comprehenders perform sensibility judgements by

using their action-planning systems to attempt to combine the sentence components into meaningful actions. Since nonsense sentences cannot be combined into coherent action plans, embodiment-type effects should, in this view, not be observed on these sentences. The results of Klatzky et al. therefore concur with action-plan based explanation of sensibility judgements.

However, other researchers have questioned what exactly, in the context of embodied cognition research, sensibility judgments tap into (Adams, 2010; Shapiro, 2010). Adams (2010) points out that a sensibility judgement cannot, itself, directly tap into comprehension; in order to determine that a sentence such as *I am hanging the coat on the teacup* is not sensible, comprehenders must first understand what it means to say *hang a coat on a teacup*. In other words, the time taken to perform a sensibility judgement is not merely the time taken to understand a sentence, but the time taken to understand the sentence plus the time taken to compare the meaning of that sentence with a set of plausible real world scenarios. As Adams notes, there may be contexts in which a sentence such as *I am hanging the coat on the teacup* can indeed be judged sensible – for example, in the context of a fashionable apartment in which the interior designer has stuck teacups to the walls to act as coat hooks.

In the absence of such specialised contexts, however, sensibility judgements rely on canonical affordances (Borghi & Riggio, 2009; Borghi, 2013). In chapters 5 and 6, I explored the assumption that future actions are usually planned according to canonical affordances, and that only in cases where the current motor context (e.g., hand-in-use) directly contradicts these affordances are they overridden. In the present chapter, I will test for a functional role of motor resonance, by manipulating the congruency between canonical affordance (i.e., the handshape typically adopted for a described manual action), and current motor context (i.e., the handshape that a comprehender is currently planning to execute). A functional account of motor resonance predicts that when the current motor context (e.g., POKE handshape) is incongruent with the canonical affordance of the sentence (e.g., PALM

handshape), sentence processing will be slower than when current motor context is congruent with the canonical affordance of the sentence.

7.3. Experiment 7

In Experiment 7, I tested whether planning a particular action causally affected sentence processing. I did this by testing whether participants who were planning to perform an action requiring a particular handshape (POKE, PALM) while reading manual action sentences, would be faster to perform sensibility judgements on sentences that described an action that was congruent with the planned action, compared with sentences that described an incongruent action. As well as manipulating handshape, I also manipulated the hand (left, right) that participants used to perform that handshape. One of the aims of this chapter was to investigate the degree of specificity captured by an action simulation. Manipulating both hand and handshape means we can distinguish between four possibilities: (1) action simulations code for affordances at both the level of handshape, and of hand; (2) action simulations code for affordances at the level of handshape, but not at the level of hand; (3) action simulations code for affordances at the level of hand, but not at the level of handshape; (4) action simulations do not code for affordances at the level of either hand, or handshape.

This experiment will extend the findings of Chapters 5 and 6 in two ways. First, it will search for evidence of a functional role of motor resonance on online language processing. Second, it will establish whether information required to perform specific actions is captured by an embodied agent's perspective. Both Klatzky et al. (1989), and Aravena et al. (2010) found effects of congruency for implied handshape, although both studies measured this congruency with a motor response. If the relationship between action and language is bi-directional, then we should also find an effect of congruency between planned action and implied handshape on VOTs in the present study. Such an effect of congruency would also suggest that the semantics of an action are specified at the level of

canonical affordances (i.e., handshape) *for a particular action*. A main effect of planned action on VOTs, but no interaction with implied handshape, would indicate a more *general effect* of action on language, but that does not code for the canonical affordances of a particular action. An effect of hand on VOTs would indicate that participants were producing action simulations based on long-term motor experience. Finally, an interaction between hand and congruency would indicate that the semantics of an event are specified at the level of both dominant hand and the handshape required for a particular action.

7.3.1. Method

Participants

Twenty-four native speakers of British English (17 female; mean age = 22 years) took part in return for course credit or payment. All participants had normal or corrected to normal vision, had no history of reading difficulties, and were right-handed (mean EHI score = 79, EHI score range = 43 to 100).

Materials and design

We constructed 80 experimental sentences, each describing a first-person manual action. Forty POKE sentences described an action typically performed with a “poking” handshape, in which the index finger is extended (e.g. *I am ringing the doorbell*), and forty PALM sentences described an action typically performed with a “flat palm” handshape (e.g., *I am giving a high five*). In addition, we constructed eighty filler, non-sensible sentences in which the verb-noun pairing was implausible (e.g., *I am folding the teapot*). All verbs in the filler sentences could, if paired with more plausible objects, denote manual actions. A total of four photographs were also prepared. Two photographs (one left hand, one right hand) showed a hand performing a POKE handshape by pressing number 5 on a number pad, and two photographs (one left hand, one right hand) showed a hand performing a PALM handshape by pressing all the keys on a number pad. Photographs showing a right hand

required a right hand response, and photographs showing a left hand required a left hand response. Given the null effect of image orientation in Chapters 5 and 6, and the fact that all sentences in this experiment were presented in the first-person, all photographs were taken from an internal orientation.

A critical item consisted of a critical (sensible) sentence plus one of the four photographs. A filler item consisted of a filler (nonsense) sentence plus one of the four photographs. Critical trials could be either CONGRUENT or INCONGRUENT. In the CONGRUENT condition, the handshape implied by the sentence (POKE, PALM) matched the handshape shown in the photograph (POKE, PALM). In the INCONGRUENT condition, the handshape implied by the sentence did not match the handshape shown in the photograph. Congruency (CONGRUENT, INCONGRUENT) was fully crossed with hand (RIGHT, LEFT). Sentences rotated through each of these four possibilities in a Latin Square design, so that each participant saw 20 sentences per condition, and every sentence was seen by an equal number of participants in each condition.

Pre-testing was carried out to minimise the possibility of semantic priming between POKE and PALM handshapes in the photographs, and the verbs in the critical sentences. Twelve native speakers of British English were shown the four photographs, and asked to describe the action they saw. Participants described the POKE images using the verbs *poke*, *point*, *prod*, *press*, and *tap*. Participants described the PALM images using the verbs *slap*, *wave*, *chop*. Therefore, the critical sentences were constructed to avoid using these verbs. In addition, POKE sentences avoided the word “finger”, and PALM sentences avoided the word “palm”. A second round of pre-testing was carried out to check that people did indeed associate POKE sentences with the POKE handshape and PALM sentences with the PALM handshape. A new sample of twelve native British English speakers were given a list of all experimental sentences, interspersed with 80 other sentences describing manual actions, and were asked to mime each action. All participants mimed a POKE action for POKE sentences, and a PALM action for PALM sentences.

Procedure

Participants made spoken sensibility judgements about sentences describing manual actions, while planning a gesture that either matched or mismatched the handshape typically associated with action described in the sentence. Participants sat at a computer terminal with a viewing distance of 60 cm. At the start of each trial a central fixation cross appeared on screen (1000 ms). A photograph appeared in the centre of the screen (1000 ms), followed by a blank screen (500 ms) and then a sentence. All text was presented in 24 point black courier new font, on a white background. Participants were instructed to silently read the sentence, and decide whether it was sensible or not. Participants said *good* out loud to indicate a sensible sentence, and *bad* to indicate a non-sensible sentence. Participants were instructed to respond as quickly and accurately as possible. Following the sensibility judgement, a blank blue screen cued participants to perform the action they had seen in the photograph by using the appropriate hand to press either number 5 (POKE), or all buttons (PALM) on a number pad (see Figure 7-1). If the participant failed to make a spoken sensibility judgement within 3000 ms of sentence presentation, the sentence disappeared and the participant moved onto the blue action cue screen. If the participant failed to register a key press on the number pad within 3000 ms of the cue screen appearing, the trial timed out and the participant moved on to the next trial. Testing was recorded using a zoom Q3 video camera, so that participant's hand responses could be coded as correct or incorrect. Voice Onset Times (VOTs) were recorded using the sound key on an SR response box. Auditory responses were recorded separately using a USB microphone to allow the experimenter to code participant's spoken responses as correct or incorrect. Button press RTs were also recorded from the number pad. The order of sentence presentation was randomised for each participant. The main experimental session was preceded by a practice session of 12 trials.

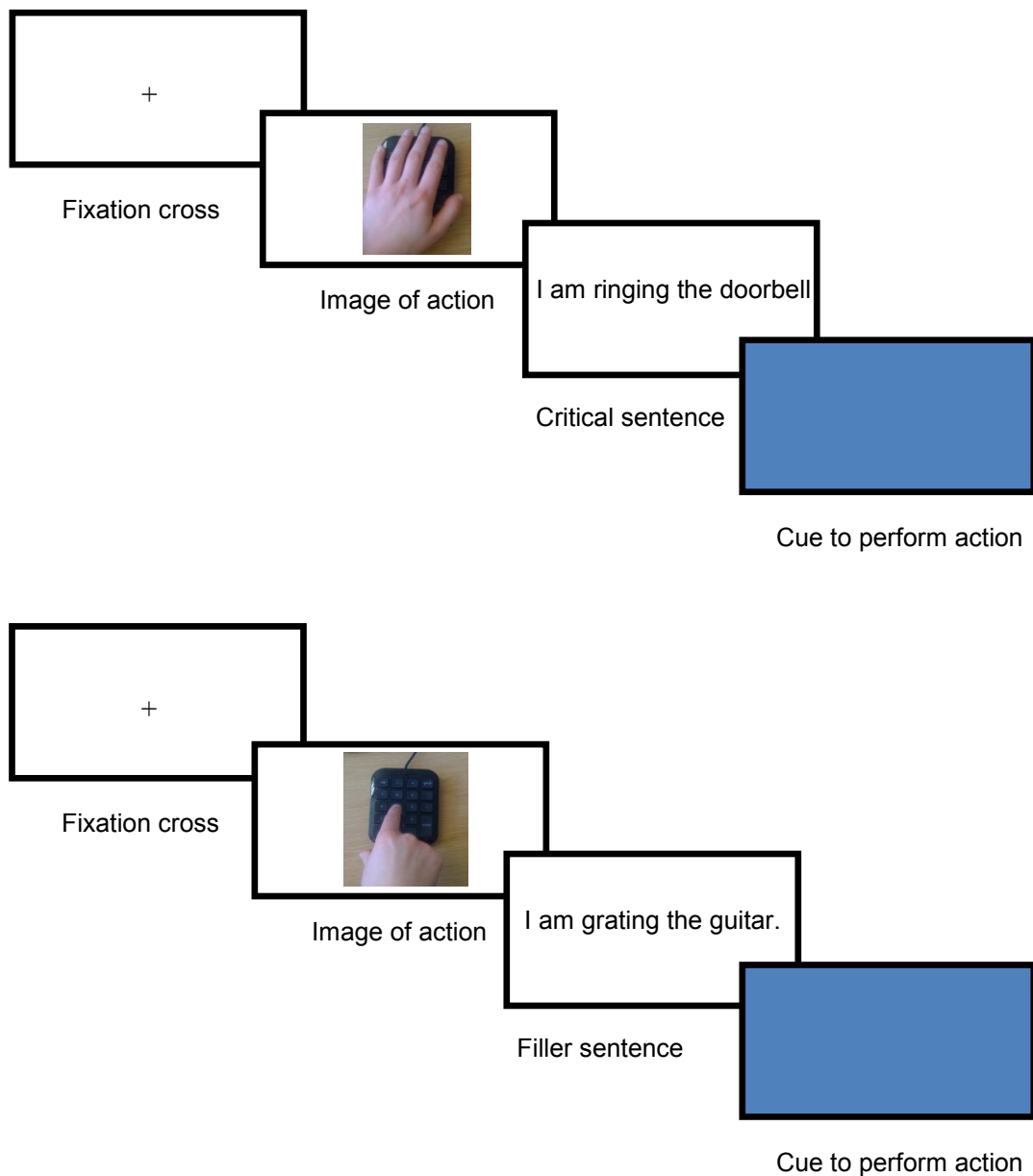


Figure 7-1. Example stimuli in critical and filler conditions in Experiment 7

7.3.2. Results

Analysis

I analysed VOTs and button press RTs from correct critical (sensible) trials, and error rates (sensibility judgment correct, performed action correct) for critical trials. Predictors of interest were sentence (i.e., the handshape typically associated with the

described action; POKE, PALM), action (i.e., the handshape participants were planning to make while reading the sentence; POKE, PALM) and response hand (left hand, right hand). An interaction between sentence and action on VOTs would constitute evidence of a functional role of motor resonance (based on action planning) in language processing. Unless otherwise specified, all analyses reported below use the following model with simplified random effect structure (by-items intercepts and slopes for response hand were removed to allow the model to converge):

Model 7-1: Sentence * Action * Response hand + (1 + Sentence * Action * Response hand | Subject) + (1 + Sentence * Action | Item).

Error rates

Sensibility judgements

Accuracy on spoken sensibility judgements was 95%. There no significant difference in accuracy between critical (95%) and filler trials (95%). Accuracy rates within the critical trials are shown in Table 7-1. There were no significant effects of response hand, sentence, planned action, or congruency (all $ps > .05$). Accuracy was over 90% for all participants, and all items.

Table 7-1. Accuracy rates for sensibility judgments in Experiment 7

| Response hand | | Sentence | | Planned action | | Congruency | |
|---------------|-------|----------|------|----------------|------|------------|-------------|
| Left | Right | POKE | PALM | POKE | PALM | Congruent | Incongruent |
| 95% | 96% | 95% | 96% | 95% | 96% | 95% | 96% |

Action production

Overall accuracy on action production (meaning that participants performed the correct handshape with the correct response hand) was 97% across critical and filler trials.

Participants were significantly more likely to correctly produce an action following a critical trial (98%) than a filler trial (97%; $B = -.36$, $SE = 0.18$, $p = .04$). Within the critical trials, participants were significantly more likely to correctly produce a POKE action (99%) than a PALM action (97%; $B = -0.81$, $SE = 0.34$, $p = .02$). There was no significant difference in accuracy rates for POKE sentences (98%) versus PALM sentences (98%), for right-handed (98%) versus left-handed (98%) responses, and for congruent (98%) versus incongruent (98%) trials (all $ps > .05$). Accuracy was over 90% for all participants, and all items.

Voice onset times

Analyses were carried out on correct (sensitivity judgement and action production) critical trials. I removed 18 trials (0.6%) where the participant corrected their response, 2 trials (0.1%) where the participant used dysfluencies (e.g., *erm*), and 5 trials (0.2%) where the participant repeated their answer (total = 25 trials; 0.9% of critical trials). In addition, 4.2% of critical trials timed out without any VOT response, and were excluded from analyses. Responses were checked to ensure that no response was made faster than 400 ms. The responses were then winsorised so that all responses above or below 2.5 sd for that participant were replaced with the upper or lower cut off value for that participant (0.04% responses replaced with the lower cut off value, 2% responses replaced with the upper cut off value). Table 7-2 shows the mean winsorised VOTs by condition. VOTs were analysed using Model 7-1 (see Table 7-3). There was no significant effect of sentence, planned action, or hand-in-use, and no interactions between these predictors (all $ps > .05$).

Table 7-2. Mean winsorised VOTs (ms) by condition in Experiment 7 (sd in parentheses). Congruent conditions are shown in bold.

| | Sentence POKE | Sentence PALM |
|-------------------|-------------------|-------------------|
| <i>Right hand</i> | | |
| Action POKE | 1557 (396) | 1578 (415) |
| Action PALM | 1574 (433) | 1602 (415) |
| <i>Left hand</i> | | |
| Action POKE | 1626 (405) | 1560 (411) |
| Action PALM | 1536 (409) | 1502 (390) |

Table 7-3. Predictors of VOTs by condition in Experiment 7: Coefficients from Model 7-1.

| Predictor | Coefficient | Standard Error | <i>p</i> (coefficient = 0) |
|--------------------------|-------------|----------------|----------------------------|
| (Intercept) | 1576.16 | 59.26 | <.001 *** |
| Sentence | 1.16 | 25.55 | 0.96 |
| Action | -24.91 | 25.54 | 0.33 |
| Hand | 19.38 | 25.97 | 0.46 |
| Sentence x Action | 25.28 | 51.08 | 0.62 |
| Sentence x Hand | 75.26 | 51.03 | 0.14 |
| Action x Hand | 88.42 | 51.04 | 0.08 |
| Sentence x Action x Hand | -21.33 | 102.07 | 0.83 |

Button press RTs

I removed trials with button press RTs of under 200 ms (1% critical trials), and a further 0.5% of trials that timed out without a response. I winsorised the remaining correct critical trials so that RTs above or below 2.5 sd for a participant's mean response latency were replaced with the upper or lower cut off value for that participant (0.18% trials replaced with lower cut off; 2.5% trials replaced with upper cut off). Table 7-4 shows the mean winsorised button press RTs by condition. I analysed the button press RTs using Model 7-1

(see Table 7-5). Participants were significantly faster to perform POKE actions compared with PALM actions. There was no significant effect of sentence or of hand (all $ps > .05$); however, the interaction between these two predictors was significant (right hand + POKE sentence = 661 ms; right hand + PALM sentence = 679 ms; left hand + POKE sentence = 673 ms; left hand + PALM sentence = 664 ms). There was no effect of congruency; the interaction between planned action and sentence was not significant ($p > .05$).

Table 7-4. Mean winsorised button press RTs (ms) by condition in Experiment 7 (sd in parentheses). Congruent conditions are shown in bold.

| | Sentence POKE | Sentence PALM |
|-------------------|------------------|------------------|
| <i>Right hand</i> | | |
| Action POKE | 626 (172) | 640 (185) |
| Action PALM | 696 (214) | 717 (241) |
| <i>Left hand</i> | | |
| Action POKE | 644 (194) | 633 (176) |
| Action PALM | 702 (220) | 694 (205) |

Table 7-5. Predictors of button press RTs by condition in Experiment 7: Coefficients from Model 7-1. Significant fixed effects are shown in bold.

| Predictor | Coefficient | Standard Error | p (coefficient = 0) |
|--------------------------|--------------|----------------|-----------------------|
| (Intercept) | 667.26 | 24.60 | <.001 *** |
| Sentence | 4.21 | 6.80 | 0.53 |
| Action | 65.83 | 16.72 | <.001 *** |
| Hand | -0.11 | 8.36 | 0.99 |
| Sentence x Action | 4.96 | 14.26 | 0.73 |
| Sentence x Hand | 27.66 | 13.59 | 0.04 * |
| Action x Hand | 13.15 | 13.60 | 0.34 |
| Sentence x Action x Hand | 2.09 | 27.18 | 0.94 |

7.3.3. Discussion

In Experiment 7, I found no effect of congruency between planned action (the handshape that participants were planning to make), and sentence (the handshape implied in the sentence), on how quickly participants performed a spoken sensibility task on sentences describing manual action sentences. Neither was there any main effect of planned action, of sentence, or of the hand that participants were planning to use. These results suggest that action simulations do not code for affordances at the level of either hand, or handshape; in fact, they provide no evidence for the existence of action simulations in language processing.

In contrast to the absence of a significant effect of action on language, analysis of button press RTs indicated an effect of language on action. Although there was no main effect of hand on button press RT, there was an interaction between sentence and hand: participants were faster to respond to POKE sentences when planning to use their right hand, but faster to respond to PALM sentences when planning to use their left hand. These results suggest that motor prototypes affect motor response (i.e., right-handed participants would typically press buttons using a poke action on their right hand). In addition, there was a main effect of planned action: participants were faster to reproduce POKE actions compared with PALM actions. This result may also reflect that in general, people tend to interact with number pads using a single finger button press.

One possible explanation for the absence of any effect of congruency on VOTs, is the difficulty of the task. If processing resources are allocated elsewhere, this may lead to more shallow processing of the sentence, and thus less detailed action simulations. Such an explanation is consistent with the view that comprehension is a fault-tolerant process (Taylor & Zwaan, 2009, 2013), and also with the view that motor activation acts as a form of priming in language comprehension (section 2.3.5), but not with strong embodiment accounts in which meaning constitutes action representations. Resource allocation might explain why other studies have shown embodiment type effects during blocking but not

randomisation of response effector (e.g., Borghi & Scorolli, 2009); it could be the case that this procedure relieves some of the burden of the task and thus affords comprehenders a deeper level of processing. In addition, several participants in Experiment 7 reported during debriefing that they had difficulty discriminating the left or right hands (although this difficulty was not reflected in error rates, with all participants achieving 90% or above in action production). Judging whether a visual stimulus represents a left or a right hand appears to involve some form of mental simulation (Parsons, 1994; Parsons, 1987); it is conceivable that this simulation might then interfere with simulation of action sentence itself. For these reasons, it seemed advisable to create a less cognitively demanding version of Experiment 7, in which response hand was blocked.

7.4. Experiment 8

Experiment 8 was a repeat of Experiment 7, except that response hand was blocked rather than randomised. This blocking manipulation reduced the processing demands of the task in two ways. First, by removing one aspect of memory (left or right hand) that participants needed to keep in mind throughout sentence processing, and second, by removing any possible effect of mental rotation in order to determine the handedness of the stimulus hand. In this way, it was hoped that participants would have more processing resources available to focus on the handshake they planned to perform, and on the sentence processing itself. As in Experiment 7, a fine-grained, action-specific functional account of motor resonance would predict an effect of congruency (interaction between sentence and planned action) on VOTs. A more coarse-grained, generalist functional account would predict an effect of planned action on VOTs, but no interaction with sentence.

7.4.1. Method

Participants

Twenty-four native speakers of British English (18 female; mean age = 21 years) took part in return for course credit or payment. All participants had normal or corrected to normal vision, had no history of reading difficulties, and were right-handed (mean EHI score = 70, EHI score range = 50 to 100).

Materials and design

Materials were the same as Experiment 7. Design was the same as Experiment 7, except that left- and right-handed images (and therefore response hand) were now blocked. Whether participants saw right- or left-handed images first was counterbalanced across participants and lists.

Procedure

Procedure was the same as Experiment 7.

7.4.2. Results

Analysis

Analysis of error rates, VOTs and button press RTs was the same as Experiment 7, using Model 7-1.

Error rates

Sensibility judgements

Accuracy on spoken sensibility judgements was 93%. There no significant difference in accuracy between critical (93%) and filler trials (94%). Accuracy rates within the critical trials are given in Table 7-6. There were no significant effects of response hand,

sentence, planned action, or congruency (all $ps > .05$). Accuracy was over 90% for all participants, and all items.

Table 7-6. Accuracy rates for sensibility judgments in Experiment 7

| Response hand | | Sentence | | Planned action | | Congruency | |
|---------------|-------|----------|------|----------------|------|------------|-------------|
| Left | Right | POKE | PALM | POKE | PALM | Congruent | Incongruent |
| 93% | 94% | 94% | 93% | 93% | 94% | 93% | 93% |

Action production

Overall accuracy on action production (meaning that participants performed the correct handshake with the correct response hand) was 94% across critical and filler trials. There was no significant difference in accuracy between critical trials (93%) and filler trials (94%). Within the critical trials, participants were significantly more likely to correctly produce a POKE action (96%) than a PALM action (91%; $B = -0.87$, $SE = 0.26$, $p < .001$). There was no significant difference in accuracy rates for POKE sentences (93%) versus PALM sentences (93%), for right-handed (93%) versus left-handed (94%) responses, and for congruent (94%) versus incongruent (93%) trials (all $ps > .05$). Accuracy was over 90% for all participants, and all items.

Voice onset times

Analyses were carried out on correct (sensibility judgement and action production) critical trials. I removed 20 trials (0.62%) where the participant corrected their response, 2 trials (0.13%) where the participant used dysfluencies, and 3 trials (0.16%) where the participant repeated their answer (total = 25 trials; 0.91% of critical trials). In addition, 2%

of critical trials timed out without any VOT response, and were excluded from analyses. Responses were checked to ensure that no response was made faster than 400 ms. The responses were then winsorised so that all responses above or below 2.5 sd for that participant were replaced with the upper or lower cut off value for that participant (0% responses replaced with the lower cut off value, 1.8% responses replaced with the upper cut off value). Table 7-7 shows the mean winsorised VOTs by condition.

I analysed VOTs using Model 7-1 (see Table 7-8). As shown in Table 7-8, there was no significant effect of congruency: the interaction between sentence handshape and planned action was not significant. There was no significant main effect of sentence handshape, or hand, and no other interactions between hand, planned action, and sentence handshape were significant (all $ps > .05$). There was a main effect of planned action: participants were faster to respond to sentences when they were planning to make a POKE action, compared with a PALM action.

Table 7-7. Mean winsorised VOTs (ms) by condition in Experiment 8 (sd in parentheses). Congruent conditions are shown in bold.

| | Sentence POKE | Sentence PALM |
|-------------------|-------------------|-------------------|
| <i>Right hand</i> | | |
| Action POKE | 1598 (422) | 1553 (417) |
| Action PALM | 1605 (430) | 1629 (462) |
| <i>Left hand</i> | | |
| Action POKE | 1557 (403) | 1563 (406) |
| Action PALM | 1583 (414) | 1599 (458) |

Table 7-8. Predictors for VOTs in Experiment 8: Coefficients from Model 7-2.
Significant fixed effects are shown in bold.

| Predictor | Coefficient | Standard Error | <i>p</i> (coefficient = 0) |
|---------------------------------|--------------|----------------|----------------------------|
| <i>Voice onset times</i> | | | |
| (Intercept) | 1600.51 | 63.71 | <.001 *** |
| Sentence | -1.72 | 24.31 | 0.94 |
| Action | 29.88 | 13.78 | 0.03 * |
| Hand | 25.21 | 20.76 | 0.23 |
| Sentence x Action | 27.66 | 30.21 | 0.32 |
| Sentence x Hand | -7.89 | 24.20 | 0.74 |
| Action x Hand | 20.55 | 28.17 | 0.47 |
| Sentence x Action x Hand | 31.22 | 49.56 | 0.53 |

Button press RTs

I checked that no trials had RTs of under 200 ms, and removed the 0.8% of trials which had timed out without a response. I winsorised the remaining correct critical trials so that RTs above or below 2.5 sd for a participant's mean response latency were replaced with the upper or lower cut off value for that participant (0.11% trials replaced with lower cut off; 2.3% trials replaced with upper cut off). Table 7-9 shows the mean winsorised button press RTs by condition. Button press RTs were analysed using Model 7-1 (see Table 7-10). As these coefficients show, there was no significant effect of congruency: the interaction between sentence and planned action was not significant. There was no significant effect of hand, sentence, and no significant interactions between these variables (all *ps* > .05). The interaction between hand and planned action was marginally significant (*B* = -25.04, *SE* = 13.55, *p* = 0.06). There was a significant effect of planned action: participants were quicker to produce POKE actions compared with PALM actions.

Table 7-9. Mean winsorised button press RTs (ms) by condition in Experiment 8 (sd in parentheses). Congruent conditions are shown in bold.

| | Sentence POKE | Sentence PALM |
|-------------------|------------------|------------------|
| <i>Right hand</i> | | |
| Action POKE | 686 (208) | 690 (208) |
| Action PALM | 747 (233) | 728 (226) |
| <i>Left hand</i> | | |
| Action POKE | 688 (218) | 710 (238) |
| Action PALM | 778 (242) | 776 (240) |

Table 7-10. Predictors for button press RTs in Experiment 8: Coefficients for Model 7-2. Significant fixed effects are shown in bold.

| Predictor | Coefficient | Standard Error | <i>p</i> (coefficient = 0) |
|--------------------------|--------------|----------------|----------------------------|
| Button press RTs | | | |
| (Intercept) | 725.68 | 28.23 | <.001 *** |
| Sentence | 1.96 | 8.41 | 0.82 |
| Action | 63.94 | 11.61 | <.001 *** |
| Hand | -22.69 | 23.53 | 0.34 |
| Sentence x Action | -20.73 | 15.17 | 0.17 |
| Sentence x Hand | -14.56 | 13.53 | 0.28 |
| Action x Hand | -25.04 | 13.55 | 0.06 |
| Sentence x Action x Hand | 4.62 | 27.09 | 0.87 |

7.4.3. Discussion

In Experiment 8, in which response hand was blocked to reduce processing demands on the task, I found no effect of congruency between action (the handshape that

participants were planning to make), and sentence (the handshape implied in the sentence), on how quickly participants performed a spoken sensibility task on sentences describing manual action sentences. There was, however, a main effect of action: participants were faster to make sensibility judgements when planning POKE actions compared with PALM actions.

The fact that planned action had an effect on VOT raises the possibility that there may be some effect of motor resonance on online language processing, but only in the sense of a general effect of activating the motor system – specifically, planning a typical action (pressing a single button on response pad with index finger) results in faster sentence processing than planning an atypical action (pressing an entire number pad with flat palm). This general effect would support accounts in the literature of general effects of language on action (Creem & Proffitt, 2001; Rueschemeyer, Lindemann, et al., 2010), but falls short of the specificity predicted by embodied accounts that posit action simulations. The results suggest that although activation of the motor system may impact language processing, it does not do so in a way that interacts with the nature of the action described in that language. In other words, adopting an embodied agent's perspective might lead to a comprehender simulating performing some action, but that the simulation is not detailed enough to specify the nature of the action being simulated.

For button press RTs, I again found that participants were faster to reproduce POKE actions than PALM actions, suggesting that participants are faster to produce typical versus atypical actions; but this does not interact with hand. Unlike Experiment 7, there was no interaction between hand and sentence.

7.5. Combined analyses from Experiments 7 and 8

To ensure that no effects had been missed, a further analysis was conducted on the data from Experiments 7 and 8, by pooling the data from these two experiments and rerunning the above analyses on the combined data. So far, slightly different results have

emerged for Experiments 7 and 8 (i.e., effect of action on VOTs in Experiment 8 but not 7; interaction between hand and sentence in Experiment 7 but not 8). The analysis was therefore rerun over the two datasets to determine whether these effects would remain significant across a larger number of trials. Responses to critical trials on Experiments 7 and 8 were pooled and analysed for accuracy and for effects on voice onset times and button press RTs.

7.5.1. Results

Analysis

Unless otherwise stated, all analyses used the following model with simplified random effect structure (where the correlation parameter was removed in order to allow the model to converge):

Model 7-2: Sentence * Action * Response hand + (1| Subject) + (1| Item) + (0 + Sentence * Action * Response hand | Subject) + (0 + Sentence * Action * Response hand | Item).

Error rates

Sensibility judgements

There were no significant effects on accuracy of response hand, sentence, action, experiment, and no interactions between these factors (all p s > .05). Accuracy was over 90% for all participants, and all items.

Action production

Participants were more likely to correctly produce POKE actions (97% accuracy) than PALM actions (94% accuracy), and this difference was significant ($B = -1.10$, $SE = 0.28$, $p < .001$). There were no significant effects of response hand, sentence, and no

interactions between these factors (all $ps > .05$). Accuracy was over 90% for all participants, and all items.

Voice onset times

Analyses on trimmed data using Model 7-2 showed a significant effect of action on VOTs: participants faster to respond to sentences while planning a POKE action compared with a PALM action ($B = 21.24$, $SE = 9.77$, $p = .03$). There were no significant effects of response hand, sentence, and no interactions between these factors (all $ps > .05$).

Button press RTs

Analyses on trimmed data using Model 7-2 showed a significant effect of action on button press RTs, with participants faster to perform POKE actions compared with PALM actions ($B = 68.44$, $SE = 10.46$, $p < .001$). There were no significant effects of response hand, sentence, and no interactions between these factors (all $ps > .05$).

7.5.2. Discussion

Across the two experiments, there was a significant effect of action on voice onset time and on button press RT: participants were faster to perform spoken sensibility judgments while planning to perform a POKE compared with a PALM action; and were also faster to execute POKE compared with PALM actions. These results suggest there may be a general effect of the motor system on action language processing – preparing an action in line with canonical affordances (e.g., pressing a single button on a number pad) results in faster sensibility judgements than preparing an unusual action (e.g., pressing a number pad with a flat palm). However, this pattern of results could also be explained by an effect of task difficulty. Planning a canonical action is, presumably, less demanding than planning an unusual action, in which no motor prototype is available (see Borghi & Riggio, 2009). It may be that sensibility judgments were facilitated by the simpler task of planning POKE actions, rather than by greater motor activation triggered by the canonical nature of POKE

actions. The current data are insufficient to distinguish between these possibilities. At any rate, it is clear that there was no evidence for an effect of action on language that codes for specificity of action implied in the sentence (either in terms of the handshape, or the hand carrying out the action).

7.6. Experiment 9

In Experiments 7 and 8, participants were instructed to perform a sensibility judgement; the congruency of the handshape implied in the sentence, and the handshape the participant was planning to make was irrelevant to the task. These instructions may have encouraged shallower processing of the sentences, such that participants only processed the sentences sufficiently deeply to form good enough (Ferreira, Bailey, & Ferraro, 2002) representations to complete the task at hand. In Experiment 9 therefore, I explicitly asked participants to judge whether the handshape in the sentence (POKE, PALM) matched the handshape they were planning to make (POKE, PALM). Studies have shown that embodied effects vary depending task demands (see section 2.4); importantly, some research suggests that when a semantic dimension (e.g., semantic size) is irrelevant to task goals, no effects are obtained, but when it is relevant, effects are found (Hoedemaker & Gordon, 2013; but see Pulvermüller, Hauk, Nikulin, & Ilmoniemi, 2005).

In addition, explicitly asking participants to make a congruency judgement makes it possible to verify to what extent participants are matching sentence handshape with planned action in the congruent versus incongruent conditions. Only materials which achieved 100% consistent responses in pre-testing were included in the experiments, but it is possible that under the time constraint of the task, certain sentences are not clearly associated with one action over another. If participants are systematically not categorising congruent sentences as congruent, this may explain the null effects of congruency on VOTs in Experiments 7 and 8. Note that, since the task in Experiment 9 is an explicit congruency judgement, affirmative (*good*) responses will overlap entirely with congruent trials. This is in contrast to

Experiments 7 and 8, when a sentence might be judged as sensible (*good*) on either a congruent or an incongruent trial. Therefore, in order to demonstrate a functional role of motor resonance on online sentence processing in Experiment 9, we must also show either a main effect of hand or an interaction between hand and congruency, in order to rule out a simple advantage for affirmative responses.

7.6.1. Method

Participants

Twenty-four native speakers of British English (11 female; mean age = 23 years) took part in return for course credit or payment. All participants had normal or corrected to normal vision, had no history of reading difficulties, and were right-handed (mean EHI score = 79, EHI score range = 45 to 100).

Materials and design

Materials and design were the same as Experiments 7 and 8, with the exception that filler trials were discarded so that only sensible sentences remained. So that the overall length of the experiment remained consistent with Experiments 7 and 8, the experiment was repeated, with each sentence appearing once in each half of the experiment, once in the congruent condition, and once in the incongruent condition. Each half of the experiment contained equal numbers of congruent and incongruent trials. Design was the same as Experiment 7, with left- and right-handed images (and therefore response hand) randomised once again.

Procedure

Procedure was the same as Experiment 7 and 8, with the exception that participants were instructed to respond whether the handshape implied by the sentence (POKE, PALM) matched the handshape in the image (POKE, PALM). Participants responded by saying

good out loud if the handshapes in the sentence and image matched, and *bad* if they did not match.

7.6.2. Results

Analysis

Analyses were similar to Experiments 7 and 8, except that because participants' task was an explicit congruency judgement, the interaction between sentence handshape and image handshape represented not only a congruency effect, but also an effect of response (*good* or *bad*). Unless otherwise stated, all analyses used the following model, with maximal random effect structure:

Model 7-3: Sentence * Action * Response hand + (1 + Sentence * Action * Response hand | Subject) + (1 + Sentence * Action * Response hand | Item).

Error rates

Congruency judgements

Accuracy on spoken congruency was 92%. Two items had accuracy 50% or lower (*I am admiring my new engagement ring; I am inspecting the mantelpiece for dust*), and were removed from all further analyses. All remaining items had accuracy of 80% or above, and all participants had an accuracy rate of 85% or over. Accuracy rates are shown in Table 7-11. Participants showed a weak tendency to make more correct congruency judgements on congruent trials (93%) than on incongruent trials (90%), although this difference was only marginally significant ($B = 0.37$, $SE = 0.23$, $p = .09$). Participants were more likely to respond correctly to sentences while planning an action with their right hand (94%) than their left hand (92%), and this difference was significant ($B = 0.24$, $SE = 0.11$, $p = .02$). There were no significant effects of sentence or of action, and no significant interactions between these factors (all $ps > .05$).

Table 7-11. Accuracy rates for sensibility judgments in Experiment 9

| Response hand | | Sentence | | Planned action | | Congruency | |
|---------------|-------|----------|------|----------------|------|------------|-------------|
| Left | Right | POKE | PALM | POKE | PALM | Congruent | Incongruent |
| 92% | 94% | 93% | 92% | 93% | 92% | 93% | 90% |

Action production

Accuracy on action production was 97%. Participants were more likely to correctly produce an action on congruent (98%) compared with incongruent trials (96%), and this difference was significant ($B = 1.14$, $SE = 0.40$, $p < .01$). There were no significant differences in accuracy between right-handed (98%) and left-handed (97%) trials, between POKE sentences (98%) and PALM sentences (97%), or between POKE actions (98%) and PALM actions (97%), and no interactions between these factors (all $ps > .05$). Accuracy was over 90% for all participants, and all items.

Voice onset times

Of critical trials, I removed 19 trials (0.63%) where the participant corrected their response, 1 trial (0.03%) where the participant used a disfluency, and 6 trials (0.20%) where the participant repeated their answer (total = 26 trials; 0.86% of critical trials). In addition, 2.1% of critical trials timed out without any VOT response, and were excluded from analyses. Analyses were carried out on correct (congruency judgement and action production) critical trials only. Responses were checked to make sure there were no responses made faster than 400 ms. The remaining responses were then winsorised so that all responses 2.5 sd above or below a participant's mean were replaced with the upper or lower cut off value for that participant (0.04% responses replaced with lower cut off; 2.19%

responses replaced with upper cut off). Table 7-12 shows the mean winsorised VOTs by condition.

VOTs were analysed using Model 7-3 (see Table 7-13). There was a significant effect of congruency: participants were slower to respond to sentences when planning a congruent than an incongruent action. There was also a main effect of image gesture, in the opposite direction to Experiments 7 and 8: participants were quicker to respond to action sentences when planning a PALM compared with a POKE action. There was no significant effect of sentence handshape, image handshape, hand, nor any further interactions between these factors (all $ps > .05$).

Table 7-12. Mean winsorised VOTs (ms) by condition in Experiment 9 (sd in parentheses). Congruent conditions are shown in bold.

| | Sentence POKE | Sentence PALM |
|-------------------|-------------------|-------------------|
| <i>Right hand</i> | | |
| Action POKE | 1442 (359) | 1501 (361) |
| Action PALM | 1464 (345) | 1403 (358) |
| <i>Left hand</i> | | |
| Action POKE | 1434 (369) | 1458 (388) |
| Action PALM | 1469 (365) | 1377 (386) |

Button press data

Analyses were carried out on correct (sensitivity judgement and action production) critical trials only. I excluded 1.3% trials where a response was made faster than 200 ms. The remaining responses were then winsorised so that all responses 2.5 sd above or below a participant's mean were replaced with the upper or lower cut off for that participant (0.09% responses replaced with lower cut off; 1.8% responses replaced with upper cut off). Table 7-12 shows the mean winsorised button press RTs by condition. Button press RTs were analysed using Model 7-3 (see Table 7-13). As these coefficients show, participants were significantly slower to produce actions on congruent than incongruent trials. There was a significant effect of action, but in the opposite direction to Experiments 7 and 8: participants were quicker to execute PALM actions than POKE actions. There was no effect of response

hand, sentence, and no further interactions between action, sentence, and hand (all $ps > .05$).

The effects on button press RTs therefore echo those found on VOTs in the same

Experiment.

Table 7-13. Predictors for VOTs and button press RTs in Experiment 9: Coefficients from Model 7-3. Significant fixed effects are shown in bold.

| Predictor | Coefficient | Standard Error | p (coefficient = 0) |
|-----------------------------|---------------|----------------|-----------------------|
| Voice onset time | | | |
| (Intercept) | 1452.82 | 38.78 | <.001 *** |
| Sentence | 19.75 | 25.62 | 0.44 |
| Action | -30.52 | 11.78 | .001 ** |
| Hand | 18.29 | 17.45 | 0.29 |
| Sentence x Action | 120.31 | 29.34 | <.001 *** |
| Sentence x Hand | -32.84 | 35.89 | 0.36 |
| Action x Hand | -19.08 | 17.33 | 0.27 |
| Sentence x Action x Hand | 23.35 | 32.47 | 0.47 |
| Button press RT | | | |
| (Intercept) | 701.45 | 55.52 | <.001 *** |
| Sentence | 9.63 | 5.26 | 0.06 |
| Action | -55.97 | 14.24 | <.001 *** |
| Hand | -1.41 | 9.69 | 0.88 |
| Sentence x Action | 87.86 | 18.03 | <.001 *** |
| Sentence x Hand | -28.47 | 17.19 | 0.10 |
| Action x Hand | 25.58 | 13.10 | 0.05 |
| Sentence x Action x Hand | -8.61 | 20.44 | 0.67 |

Table 7-14. Mean winsorised button press RTs (ms) by condition in Experiment 9 (sd in parentheses). Congruent conditions are shown in bold print.

| | Sentence POKE | Sentence PALM |
|-------------------|------------------|------------------|
| <i>Right hand</i> | | |
| Action POKE | 712 (294) | 746 (276) |
| Action PALM | 693 (271) | 672 (269) |
| <i>Left hand</i> | | |
| Action POKE | 726 (230) | 762 (279) |
| Action PALM | 716 (241) | 630 (310) |

7.6.3. Discussion

In Experiment 9, I found an effect of hand-in-use on how accurately participants performed a spoken congruency judgement: participants were more likely to correctly identify a sentence as congruent with their planned action with they were planning to perform that action with their right (dominant) hand, compared with their left (non-dominant) hand. For VOT analysis, an effect of congruency was observed, with participants faster to respond to congruent versus incongruent trials. However, there was no effect of hand nor any interaction with hand, therefore this finding might also be explained without positing any interaction between action and language. For example, it could be the result of increased processing difficulties for negative compared with affirmative concepts (e.g., Wales & Grieve, 1969) or the fact that the congruency between action and sentence may have led to an increased feeling of perceptual fluency in participants compared to incongruent conditions (Reber, Winkielman, & Schwarz, 1998). Further research, in which congruency (congruency, incongruent) could be distinguished from response (good, bad) are required to resolve this issue.

Experiment 9 also showed an effect of planned action on both sensibility judgment and action execution, however this effect was in both cases in the *opposite* direction to that found in Experiments 7 and 8. Participants in Experiment 9 were faster to respond to sentences while planning PALM actions compared with POKE actions. Both of these effects (congruency, planned action) remained when analysing only the first half of the experiment.

The reversal in effects of planned action on both VOTs and button press RTs between experiments 8 and 9 is unexpected. One possibility is that in the explicit congruency task, participants were more likely to use motor imagery to consciously visualise the actions described in the sentences. In this case, if PALM actions tended to be more clearly associated with the PALM sentences than POKE actions were with the POKE sentences, this might result in longer processing times when participants planned a POKE action. However, there is no evidence of increased difficulty associating POKE actions and sentences in the error rates (section 7.5.2).

The effect of hand on accuracy (more correct responses when planning to execute a right-handed response) suggests a possible effect of long-term motor experience on language processing. In particular, it suggests that participants might run action simulations that code for the hand typically used to perform an action, but not for more detailed information such as handshake. The fact that this effect emerged when the task was an explicit congruency judgment, but not when it was a sensibility judgment (Experiments 7 and 8), suggests that participants may only assign a handedness representation when the task requires them to pay attention to the hand implicated in the sentences. Other behavioural research also indicates that attention may influence how likely embodied effects are to emerge (Hoedemaker & Gordon, 2013; Tomasino et al., 2007). Interestingly, the Linguistic Focus hypothesis (Taylor & Zwaan, 2008; Zwaan, Taylor, & de Boer, 2010) provides evidence that motor simulations occur when the comprehender's attention is drawn to the action in the sentence, and subside once linguistic focus shifts attention to other aspects of the sentence (e.g., what the agent in the sentence is thinking). Explicit tasks such as the congruency judgment in Experiment 9 may serve a similar function as linguistic focus in directing the comprehender's attention to aspects of the sentence most relevant for action simulations (see Tomasino & Rumiati, 2013, for a recent discussion of possible top-down influences on motor activation in language). Future research should therefore aim to elucidate the role of attention in moderating referential motor resonance.

7.7. General discussion

In three experiments, I tested for evidence of a functional role for motor resonance on sentence processing. I tested whether congruency between the handshape implied in a sentence and the handshape a participant planned while reading the sentence, affected language processing. In Experiments 7 and 8, where congruency was irrelevant to the task, I found no evidence of any effect of congruency or hand-in-use on sentence processing, in either accuracy or VOTs. In Experiment 9, where participants made an explicit congruency judgement, I found an effect of response hand on accuracy, but not on VOTs.

The lack of an effect of congruency in Experiments 7 and 8 runs counter to the predictions of embodied theories that posit an action-specific causal role of motor resonance on online language comprehension. The fact that, in Experiments 8 and 9, the speed of spoken sensibility judgments was influenced by which action (POKE, PALM) participants were planning to perform while reading the sentence, hints at the possibility of a very general effect of motor activation in which the comprehender simulates that she will perform some kind of action, but this simulation does not code for either the hand or the handshape typically associated with that specific action. However, the reversal of the effect direction (faster judgments while planning POKE actions in Experiment 8; faster judgements while planning PALM actions in Experiment 9) makes it difficult to draw any firm conclusions about the effect of action on language.

7.7.1. Effects of planned action

As noted above, in two of the three experiments in this chapter, there was a significant effect of the action that participants were planning to perform on the VOT measure of sentence processing. However, the direction of this effect was inconsistent. In Experiment 8, when congruency was irrelevant to task, participants were faster to respond when planning a POKE action (presumably, a more canonical affordance for interacting with

a number pad). However, in in Experiment 9, when the task was an explicit congruency judgement, participants were faster to respond to sentences when planning a PALM action (presumably, an atypical affordance for interacting with a number pad). Interestingly, in both experiments, the same effect of planned action was found on button press RTs as on VOTs: faster responses for POKE actions in Experiment 8, and faster responses for PALM actions in Experiment 9. The fact that the direction of effect switched for both sentence processing and action execution suggests that the effects on VOTs may in fact have represented a wider effect of the task on processing in general, rather than an effect on language processing in particular. For example, in Experiment 8, when handshape was irrelevant to the task, participants may simply have found it easier to plan a canonical action (POKE) than a non-canonical action (PALM), thus facilitating downstream processing (both linguistic and action execution). In contrast, in Experiment 9, when handshape was relevant to the task, participants may have found PALM sentences were more clearly associated with PALM actions than POKE sentences were with POKE actions, thus facilitating processing on both the language and action parts of the task.

The fact that participants did not appear to simulate the specifics of a particular action is perhaps surprising given that motor resonance seems to be sensitive to distinctions between precision grips and power grips in both action observation (di Pellegrino et al., 1992; Giacomo Rizzolatti & Craighero, 2004), and in language based tasks (Maurizio Gentilucci & Gangitano, 1998; Glover et al., 2004). Note however, our study used POKE versus PALM gestures as these were judged to be maximally dissimilar to one another. It is possible that using equivalents of the precision grip (PINCH) and power grip (CLENCH), may have led to more conclusive results. Use of PINCH and CLENCH might also allow the materials to describe functional, rather than non-functional actions. Recent research suggests that function actions (in which an object is used) may result in increased motor resonance than non-functional actions (e.g., an object being moved or displaced (Bub et al., 2008; Buxbaum et al., 2006). In Experiments 7 – 9, many of the POKE sentences described

functional actions, but very few of the PALM sentences did so. Future research ought therefore to try and look for effects using functional sentences implying two different handshapes.

7.7.2. Effects of response hand

In Experiments 7 and 8, I found no effect on sentence processing of the hand participants were planning to use. However, in Experiment 9 I found participants were significantly more accurate when planning an action with their dominant hand. As outlined above, these conflicting findings suggest that participants only assign a handedness explanation when the task required them to explicitly consider the handshape implied in the sentence. In Chapter 6, I found an overall interaction between response hand and pronoun, indicating that comprehenders assign a handedness representation to language based on the hand they are currently planning to use, rather than the hand they have typically used in the past. However, further analyses revealed that the effect of pronoun, although still significant for right-handed responses, was not significant for left-handed responses. This further finding suggested that the sensitivity of embodied representations might be stronger when the current motor plan coincided with long-term experience. Similarly in Experiment 9, participants' representations may have been more sensitive to congruency when their current motor plan (response hand) matched their long-term experience (dominant hand).

However, it is important to remember that the effect of hand in the present chapter was very weak, and was restricted to accuracy data rather than VOTs. One obvious question is why there were no effects on VOT, when Chapters 5 and 6 had found consistent effects in a forced-choice sentence-picture matching task. There are several possible answers to this question. One possibility is the differences were there, and that the VOT paradigm used in the present chapter was simply not sensitive enough to pick up differences. This possibility could be easily tested by having participants perform manual sensibility judgements while planning actions with a different effector (e.g., foot). A second possibility is that no

differences were found in Experiment 9 precisely because we separated the effector used to respond (mouth), and the effector that formed the basis of our congruency manipulation (hand). In contrast, Chapters 5 and 6 employed a forced-choice sentence-picture matching task in which the type of effector (i.e., hand) implicated in the sentences was the same as that used to respond. The experiments reported in Chapters 5 and 6 may, therefore, have shown an effect of language on action, rather than the other way around. It is also possible that in studies in which a motor response is required (Glenberg & Kaschak, 2002; Zwaan & Taylor, 2006), the findings are in fact the result of a two-way interaction of language on action and also of action on language. A third possibility is that the results from Experiments 1 – 6 demonstrate an effect of action on how language is interpreted, after initial comprehension has occurred. Image selection in Experiments 2a – 6 occurred after sentence processing was complete. Although the interaction with pronoun rules out an image-only based explanation, it is possible that role of motor resonance may change over time, such that online sentence processing relies on more on heuristics, based on long-term motor experience (i.e., dominant hand) and that, only once sentence processing is complete and participants shift their attention from sentence processing to motor execution, does hand-in-use become the dominant force in how a sentence is interpreted. Thus, Experiments 1 – 6 may have shown offline effects of action on language. The fourth possibility is that the results from Experiments 1 – 6 are due to an explicit strategy of the participants. However, this possibility is unlikely for the reasons given in section 6.6.4, namely: the absence of any reported strategy during debriefing; the interaction with pronoun in Experiments 2a, 2b, 5, and 6; and the sudden null effect of response hand and pronoun in Experiment 4.

7.8. Summary

In this chapter, I failed to find evidence of a functional role for action-planning based motor resonance on online language processing. Specifically, there was no effect of whether, while reading a manual action sentence, participants were planning a congruent

action or not. There was some evidence suggestive of a general effect of motor activation on language processing, but no evidence that this codes for individual actions. In other words, there is some rather weak evidence that adopting an embodied agent's perspective has a functional role on language processing, but no evidence that adopting an embodied agent's perspective involves simulating fine-grained information about the specific action being described. One possible avenue for future research is how the level of detail in action simulations might be moderated by attention (driven by, for example, task demands or linguistic focus) towards relevant motor features.

8. CONCLUSIONS

8.1 Thesis aims

In this thesis, I have explored the constraints governing the interaction between the language and action systems by tackling the following three questions. First, what perspective do comprehenders adopt when processing simple action sentences? Second, is this perspective grounded in motor activation from long-term motor experience, short-term motor experience, or current action plan? Third, is there any evidence that the motor system plays a functional role in online language comprehension? In the following section, I provide a summary of these results.

8.2. Summary of results

We saw in Chapter 3 that the embodiment literature often assumes that a comprehender will represent a described action as though she were carrying out that action herself (i.e., adopt an embodied agent's perspective; Barsalou, 1999, 2003, 2008); we also saw that in fact, this is not necessarily the case. A comprehender could equally adopt the perspective of an observer or, in sentences featuring an animate patient, the perspective of that patient. To investigate what perspective was adopted, I used a verification sentence-picture matching task (Experiment 1), and a two-alternative forced-choice sentence-picture matching task (Experiments 2a – 6).

Experiments 1 – 2b took advantage of the fact that left- and right-handed people typically perform actions with different hands, but (presumably) both observe other people performing a majority of right-handed actions. The two possible perspectives in our sentences (embodied agent, embodied observer) make different predictions as to how left- and right-handed participants will behave in a sentence-picture matching task. If comprehenders adopt an embodied patient's perspective, then groups of right- and left-

handed participants should each respond in a similar way to left- and right-handed images. If comprehenders adopt an embodied agent's perspective, then right- and left-handed participants should respond differently to right- and left-handed images. Experiment 1 showed an interaction between participant's dominant hand, and the handedness of the image, suggesting that comprehenders adopt an embodied agent's perspective, in line with the Body-Specificity Hypothesis (Casasanto, 2009, 2011).

In a two alternative forced-choice sentence-picture matching task, Experiment 2a found that the effect of dominant hand was stronger for first-person sentences than for third-person sentences, supporting the suggestion the self-referential language is "more" embodied, because the comprehender can ground the representation in their own body (see Chapter 3). However, Experiment 2b suggested that these results were due to response hand rather than dominant hand, and Experiment 3 found no effect of dominant hand when responses were measured using voice onset time, thus eliminating manual responses from the task. Both of these findings, especially that of Experiment 3, run counter to the Body-Specificity Hypothesis's prediction that the way in which we interpret action language is shaped by our long-term motor experience.

In Chapter 6, I investigated whether the effect of response hand in Experiment 2b was due to short-term motor experience (the hand used to respond on a previous trial) or to current motor plan (the hand the participant was planning to use while reading the sentence). I recruited right-handed participants and manipulated at what point during a forced-choice sentence-picture matching task the participant was able to form a motor plan for the correct response hand. Results replicated the response hand by pronoun interaction from Experiments 2a and 2b, but crucially, only when participants knew their response hand during sentence processing (and were thus able to form a motor plan for that hand). There was no effect of the hand used to respond on the previous trial, and no effect of dominant hand. These findings suggest that the interaction between action and language is based on motor resonance from motor planning, rather than from re-enactment of previous actions

(either accumulated long-term experience, or immediate short-term experience). One implication of these results is that motor resonance in language may have a predictive role, rather than merely simulating past experiences. The effect of pronoun in the right-handed, but not left-handed responses of right-handers in Experiments 5 and 6 suggests that action simulations may code features such as agency in more detail when motor planning and motor experience coincide.

In Chapter 7, I tested for a causal effect of planned action on language processing. Right-handed participants performed a sensibility judgement on manual action sentences, a paradigm which has consistently shown an effect of language on action (Borreggine & Kaschak, 2006; Glenberg & Kaschak, 2002; Kaschak & Borreggine, 2008). In order to test for an effect of action on language, rather than the other way round, participants were required to plan a manual action and then make a verbal response on the sensibility judgment. In other words, the dependent measure of sentence processing was not a manual action. In Experiments 7 and 8, I found no evidence to support a functional role of action planning in language processing. In Experiment 9, participants were explicitly asked to assess whether their planned action was congruent with the action described in the sentence. Results showed that in Experiment 9, participants were more likely to make a correct judgment when planning to perform an action with their right hand, compared with their left hand. These findings suggest that comprehenders assign a handedness to linguistic representations only when the task explicitly requires them to compare their linguistic representation with their motor representation.

8.3. Implications and future research

In this section, I will discuss the implications of these results for our understanding of language processing in general, and embodied cognition in particular. I will also discuss some possible directions for future research.

8.3.1. Perspective-taking in action language

The results from Experiments 1 - 6 imply that, in simple action sentences, comprehenders adopt the perspective of an embodied agent – in other words, they interpret the sentence as though they were performing that action themselves. These results are in line with two assumptions in the embodied cognition literature. First, that comprehenders do adopt an embodied perspective on action language. And second, that this perspective is most likely to be that of the embodied agent. Had participants in Experiments 1-6 *not* adopted an embodied perspective, then participants would have shown no effect of either dominant hand or the hand they were planning to use. Had participants adopted the perspective of an embodied *observer*, then they would have shown a preference for right-handed interpretations, regardless of their dominant hand or the hand they were planning to use. However, the results go further than simply confirming the wide-spread assumption that comprehenders adopt an embodied agent's perspective. The results also highlight some of the constraints under which this embodied perspective-taking operates.

Pronoun

Across 4 experiments (Experiments 2a, 2b, 5, 6), whenever an effect of response hand was found, this effect interacted with pronoun such that the effects were stronger following first-person sentences (e.g., *I am slicing the tomato*) than following third-person sentences (e.g., *He is slicing the tomato*). Other work in the literature has also found differential embodied effects for first- versus third-person sentences (Brunyé et al., 2009, 2011; Ditman et al., 2010), although the explanations these researchers give are somewhat at odds with present findings. For example, Brunyé et al. (2009) found that participants were faster to match first-person sentences to internally orientated, and third-person sentences to

externally orientated images². They argued that their results show that comprehenders adopt an embodied agent's perspective for first-person language, but an embodied observer's perspective for third-person language. However, we noted in the discussion of Chapter 5 that our pattern of results cannot be explained by participants adopting an embodied observer's perspective on third-person sentences, because the embodied observer's perspective would also lead to a preference for right-handed interpretations (since most observed actions are right-handed), whereas in fact, participants using their right-hand showed a *weaker* preference for right-handed interpretations of third-person sentences.

Rather, our results support the possibility that for third-person language, the embodied perspective is either less strongly embodied, or less likely to be adopted, compared with first-person language. Papeo et al. (2011) using TMS found increased cortical excitability on action versus non-action verbs only when those verbs were presented in the first-person; third-person action verbs showed no increased compared with third-person non-action verbs. From this finding, Papeo et al. conclude that action simulations are not automatically generated, but depend on, amongst other things, whether the implied agent of the action is self-referential or not. Our behavioural results support this interpretation. Importantly, Experiments 4- 6 suggest that the motor resonance implicated in language comprehension is based on prospective simulation (action planning) rather than retrospective simulation (reenacting previous experience), adding weight to recent proposals that motor resonance in language may have a predictive role (Pickering & Garrod, 2009; Rueschemeyer & Bekkering, 2013). We saw in section 2.2.4 that people are better at predicting their own actions than those of other people. Given this first-person advantage in prediction, it follows that if motor resonance has a predictive function – perhaps acting, as Pickering and Garrod (2009) propose, as a type of emulator, or forward model (Grush, 2004) – then it makes sense that embodied representations might be stronger for self-referential language compared with

² While writing this thesis, I became aware of several failures to replicate these findings: Beveridge, Nieuwland, Santesteban & Pickering (unpublished); Vukovic (2013, personal communication); Zwaan (2014, personal communication); but see Sato & Bergen, 2013 for a replication.

third-person language. Note that, although I have limited discussion to the contrast between first- versus-third person language, we should expect similar findings if the first-person stimuli were replaced with second-person stimuli.

Hand orientation

In work by Brunyé and colleagues (Brunyé et al., 2009, 2011; Ditman et al., 2010), the orientation of hands depicted in images (internal, external) appeared to cue whether participants adopted an embodied agent's perspective, or an embodied observer's perspective. Based on these findings, and similar effects of hand orientation in action observation (e.g., Vogt et al., 2003), we predicted that image orientation might act as a cue for agency attribution (perhaps, for example, participants would be more likely to adopt an embodied agent's perspective on internally orientated images). However, there was no such evidence of an effect of image orientation in this thesis – in none of the 5 experiments manipulating image orientation was there any effect of this manipulation, either alone or in interaction with other variables.

The lack of any effect of internal versus external orientation led to the suggestion that orientation might only act as an inconsistent cue to agency attribution, but it is unclear why this might be the case. In section 5.8.2, I hypothesised that congruency with the effector (right or left hand) took precedence over congruency with orientation (internal or external): in Experiments 2a- 6, two different effectors were presented simultaneously, compared with a single effector in studies by Brunyé and colleagues. One explanation of our null finding for orientation is therefore that agency attribution might operate on different levels of specificity, with more general information (e.g., effector) used as a cue where possible, and more finely-coded information (e.g., effector orientation) used when this coarser information is not available. In cases where different levels of information contradict one another (e.g., a right-handed participant viewing an externally orientated right hand), the more general cues

might override the more specific cues. Of course, such an account is speculative, and would require further testing.

A further explanation is that, rather than categorising effector and orientation as more or less general cues, we could categorise them as motor and visual cues, respectively. In Experiments 2a – 6, participants were required to perform a manual action to select one of two stimuli. Due to the manual response, the motor cue (*is this the same as the hand I am moving?*) may have taken precedence over the visual cue (*is this hand in the same position as my own hand?*). There is some support for a motor *versus* visual cue explanation, from research into action-recognition. Jeannerod (2003) reviewed studies in which adults are required to categorise actions as either their own actions or those of other people (Daprati et al., 1997; Sirigu, Daprati, Pradat-Diehl, Franck, & Jeannerod, 1999). He concluded that when self- and other-generated actions were the same, participants used visual cues (such as hand orientation) to determine agency. However, when self-generated actions differed from other-generated actions, participants used efferent signals about their own movement to determine agency. In Experiments 2a – 6, only one of the two hands in critical trials matched the effector that the participant was planning to move. Participants may therefore have ignored visual signals such as orientation, and focussed on motor signals (e.g., which stimulus hand matches the hand being activated).

The Theory of Event Coding

In section 2.2.2, I outlined the Theory of Event Coding (TEC; Hommel et al., 2001; Prinz, 1987) as a possible explanation of the link between action and perception. TEC accounts stress the shared code between language simulation, and what the intended outcome of an action will look like. In other words, effects such as the ACE might be explained by the perceptual consequences of an intended action (i.e., having your hand in a far position rather than a near position, for example), rather than by a motor representation of that action (activating the motor systems involved in making a far or near movement; see

Diefenbach et al., 2013). Although Diefenbach et al. explain the ACE effect in terms of TEC rather than motor resonance, they admit that their data do not allow them to distinguish between these possibilities, since the motor representation (making a far movement) and the intended action effect (having your hand in a far position) overlap completely.

One means of distinguishing between these two possible explanations (perceptual consequences versus motor resonance) might be to manipulate hand orientation (internal, external). Presumably, only internally orientated images would correspond to the expected perceptual outcome of an agent's action, since externally orientated hands cannot easily be interpreted as the viewer's own hands. In Experiments 2a – 6, participants matched manual action sentences to pairs of internally or externally orientated images. Given that externally orientated images seem less likely to capture the expected perceptual consequences of an action, accounts of action language understanding that stress those perceptual consequences (e.g., Diefenbach et al., 2013; Kaschak & Borreggine, 2008) might predict an interaction with image orientation (i.e., weaker effects on trials with externally orientated images). In contrast, accounts of action language processing that focus on motor activation (Glenberg & Gallese, 2012; Taylor & Zwaan, 2013) might predict embodied effects regardless of orientation. Over five experiments, I found no effect of image orientation on the way that participants interpreted manual action sentences. Rather, comprehenders seem to interpret the sentences by adopting an embodied agent's perspective, based on the hand for which they have a current action plan. Note that our results do not rule out the possibility that planned actions might share a common coding scheme with the perceptual consequences of those actions (e.g., Prinz, 1997). However, the results do suggest that this coding scheme does not capture fine-grained information about the particular hand (left or right) that would be used to perform an action.

8.3.2. *Spatial grounding and good enough representations*

Chapter 3 introduced the Spatial Grounding Hypothesis – the suggestion that, in order for comprehenders to adopt an embodied perspective, the motor simulation must be spatially grounded. For self-referential language, this grounding can occur via the comprehender's own body; however, for third-person language, additional information about the action (e.g., direction) may be required. This information could be presented through additional utterances integrated into a suitably detailed situation model (e.g., Taylor & Zwaan, 2009), or perhaps also through visual cues such as avatars (Gianelli et al., 2011). We might suppose that Taylor and Zwaan's (2013) construal of comprehension as a fault-tolerant process could be applied to the Spatial Grounding Hypothesis: in the absence of suitable spatial grounding, participants could adopt an unembodied perspective. This unembodied perspective would allow comprehension to occur, albeit perhaps not so rich as comprehension resulting from an embodied perspective.

If the Spatial Grounding Hypothesis is correct, then participants should have failed to adopt an embodied agent's perspective in the third-person sentences in Experiments 2a- 6, and instead performed at chance. However, although this might appear to be the case for right-handed participants in Experiment 2b (52% preference for right-handed image), it is less clear, for example, that this is the case for left-handed participants on Experiment 2a (65% preference for left-handed images). There are two possible explanations for the weaker (but still present) preference for hand-in-use based interpretations of third-person sentences. First, that participants do adopt an embodied agent's perspective on third-person sentences, but that this is less embodied than the corresponding representation for first-person sentences. Second, that comprehenders are less likely to adopt an embodied agent's perspective on third-person sentences. Current data are insufficient to distinguish between these alternatives. However, it is clear that the Spatial Grounding Hypothesis requires at least some modification, to account for the fact that participants did not always return to

chance performance on the third-person sentences, despite the absence of any apparent grounding material.

The idea that comprehension is fault-tolerant is reminiscent of good-enough representations in sentence processing (Ferreira et al., 2002; Ferreira & Patson, 2007). The concept of good-enough representations emerged from research into syntactic processing, which suggested that comprehenders often misinterpret sentences, and retain incorrect interpretations in memory once sentence processing is complete (Christianson, Hollingworth, Halliwell, & Ferreira, 2001; Ferreira, 2003). Traditional theories of sentence processing stress that comprehension results from a series of algorithms performed on amodal symbols, resulting in “*complete, detailed, and accurate representations of the linguistic input*” (Ferreira et al., 2002, p.11). Embodied accounts reject such theories on the basis of the computations over amodal symbols; however, the conception that comprehension is fault tolerant suggests that embodied theories might also reject the premise that comprehension typically provides complete and accurate representations.

The good-enough approach to language comprehension argues that language processing is shallower, and less complete than traditional theories assume; representations are merely good-enough for the task at hand (e.g., Sturt, Sanford, Stewart, & Dawydiak, 2004). In conversation, the goal may be to provide an appropriate response, and in a psycholinguistics experiment, the goal might be to execute a suitable motor movement; neither of these tasks necessarily requires participants to provide evidence that they have constructed a complete and detailed representation of the sentence in question (Ferreira & Patson, 2007). This being the case, the language system may simply provide the most detailed representation necessary to fulfill the current goal: comprehension might only need to be “*good enough for current purposes*” (Clark & Schaefer, 1987, p. 19). Thus, the good-enough account implies a good deal of flexibility in the depth of language processing, and this flexibility might be invoked in order to explain the fact that so many embodied effects depend on specific task requirements. For example, participants in Experiments 2a – 6 may

not have shown any effect of dominant hand, because they were able to perform the task without assigning any long-term handedness representation to the sentence. All the situation required was an interpretation based on the hand that participants were planning to use. In contrast, when participants were *explicitly* instructed to judge the congruency between action and sentence in Experiment 9, participants assigned a handedness representation to the sentence. It should be noted that, the effect of pronoun in Experiments 2a – 6 occurred in a task which explicitly required participants to attend to the pronoun of the sentence, in order to successfully recall the sentences. In section 3.4., I suggested that the disparity between those studies that found a difference between first- and third-person language (e.g., Papeo et al., 2011) and those which did not (e.g. Tomasino et al., 2007), might be explained in terms of whether or not the task required participants to pay attention to whether the stimuli were presented in first- or third-person. It should therefore be tested whether the effect of pronoun in Experiments 2a – 6 can be replicated in a task that does not require overt attention to pronoun; or alternatively whether, when such attention is not needed, comprehenders simply form good-enough representations which are not coded for person.

As Ferreira et al. (2002) suggested, good-enough representations are even more likely in real-life language use than in idealised laboratory conditions, due to the amount of noise in the signal (disfluencies, background noise, accents, etc.). One implication of this, is that comprehenders may be more likely to rely on predictive mechanisms to aid their understanding. Therefore, if motor resonance does have a predictive function, then comprehenders may show stronger effects of action on language outside the laboratory compared with inside. After a decade in which countless articles have demonstrated specific instances of embodied effects, some researchers are beginning to challenge the usefulness of an *embodied* versus *non-embodied* division in the literature (Willems & Francken, 2012). It should by now be clear that language and action do interact – but it should be equally clear that this interaction only occurs in certain circumstances. Future research should, of course, aim to specify what these circumstances are. But in addition to this, research should attempt

to integrate findings from the embodiment literature with existing frameworks and concepts in wider psycholinguistics. The application of good-enough processing to action-language comprehension would constitute an obvious starting point.

8.3.3. Action planning and affordances

Results from Experiments 4 – 6 suggest that the interaction between action and language is driven by prospective simulation. In other words, the embodied agent's perspective appears to be grounded in a comprehender's motor plan at the time of comprehension, rather than their motor experience (short-term or long-term). The finding that the interpretation of action sentences is affected by upcoming actions is suggestive of a predictive function for referential motor resonance. Borghi (2013) stresses the fact that simulation can involve both reenactment of previous sensorimotor experience (Barsalou, 1999), and preparation for future actions (Grush, 2004). Canonical affordances capture, according to Borghi and colleagues (Borghi & Riggio, 2009; Borghi, 2013), information about how we typically interact with an object, for use in future interactions. For example, we typically use a knife by holding the handle towards us. In other words, canonical affordances use long-term experience to predict future actions.

However, results from Experiments 1 – 6 indicate that there is more to it than this. Canonical affordances capture the way we are most likely to interact with an object, based on previous interactions; but when a non-canonical motor plan is active, the affordance adapts and captures instead the way we are most likely to interact with the object *at that point in time*. For example, based on canonical affordances, I would usually understand a sentence such as *I am slicing the tomato* as though I were slicing a tomato with my dominant hand. However, if, while processing this sentence, I planned to perform an action with my non-dominant hand, then I would understand the sentence as though I were slicing a tomato with my non-dominant hand (although the relevant action simulations might be less sensitive than when using my dominant hand). Such flexibility is, in fact, eminently plausible. After

all, an action system in which we were unable to adapt to the constantly changing motor context in which we operate would be of little use.

One possible mechanism for allowing such flexibility in affordances is attention. Research into affordances has typically used items with handles on one side (e.g., cups) to allow the object's orientation to be congruent or incongruent with a particular hand; participants tend to show an advantage when the orientation of the object is congruent with their response hand (Symes, Ellis, & Tucker, 2005; Tucker & Ellis, 1998). Such results are then usually interpreted in terms of affordances – in other words, a cup automatically generates a motor code based on its associated action (e.g., being held by its handle). However, Anderson, Yamagishi and Karavia (2002) argued that the use of asymmetric objects in such studies could have resulted in asymmetric distribution of attention across those objects, and that these differences in attention could account for the congruency effects without invoking automatically generated action codes. In support of this theory, Anderson et al. found that participants showed a left or right hand advantage for asymmetric non-objects, as well as objects such as cups: the non-objects were computer generated and were therefore not associated with any particular actions. Future research should therefore seek to investigate the role of attention on affordances described in language, as well as those depicted in visual objects. Indeed, it may be possible to integrate such work with an account of good-enough representations in embodied language (see section 8.3.2), given evidence that linguistic focus appears to influence both depth of processing (Sturt et al., 2004), and the effect of motor resonance (Taylor & Zwaan, 2008).

8.3.4. Implications for naturalistic language processing

Experiments 2a – 6 showed a robust pattern of results, in which participants tended to select an image that was congruent with their response hand, only when they had an action plan available during sentence processing; and that this preference was stronger in first- compared with third-person sentences. These results were obtained using a two-alternative

forced choice paradigm, which has been successfully used to investigate psycholinguistic phenomena (e.g. Raffray et al., 2007). I have argued that the pattern of results obtained cannot easily be attributed to demand characteristics or strategic processing on the part of the participants (sections 5.8.4, and 6.6.4, respectively). Nonetheless, forced-choice paradigms can be criticised due to the unnatural nature of the task: language comprehension is an online process, a fact which is not captured in the forced-choice paradigm, where image selection takes place *after* sentence processing is complete. The task therefore gives us no measure of the timing of particular aspects of processing. In Experiments 7-9, I used a sensibility judgement task, extensively used in the embodied literature (e.g., Glenberg & Kaschak, 2002), in order to give us an arguably online measure of sentence processing (Voice Onset Time). I found no significant effects of motor plan on this measure.

However, Zwaan and Taylor (2006) argue that sensibility judgements do not, in fact, reflect online processing. This argument is based on the fact that Zwaan & Taylor's reading-by-rotation paradigm appears to show that motor resonance is a short-lived phenomenon: the congruency effect they observed (participants were faster, in a self-paced reading paradigm, to read about actions involving clockwise hand movements while themselves performing a clockwise hand movement) was limited to the action verb, and the effect had dissipated by the end of the sentence. They therefore argue that the effects found in tasks occurring after sentence presentation is complete, such as sensibility judgements, might instead be attributed to a post-sentential re-simulation of the described action, performed in order to complete the task. If this is the case, then our task in Experiments 7-9 failed to find an effect of action plan on this second "re-simulation" stage of motor resonance, but did not tap into the first, online stage. One avenue for future research would therefore be to develop a paradigm that can investigate this posited first, localised stage of motor resonance, perhaps through the use of eye-tracking in a silent reading task. The use of eye-tracking would allow researchers to investigate, with some precision, various critical regions of interest, and thus to determine

the extent to which motor resonance is a short-lived and localised phenomenon as proposed by Zwaan and Taylor.

A second potential criticism of the tasks used in this thesis, is that stimuli were limited to descriptions of very simple, manual actions, such as slicing a tomato or ringing a doorbell. More complex actions involving numerous sub-goals over variable time periods (e.g. carrying a parcel to the post-office), or non-action descriptions using abstract or metaphorical language (e.g. understanding the theory about unicorns), were not considered. It is certainly true that embodied theories of language must provide an account of how more complex language is understood. However, the relationship between motor activation and language comprehension has not yet been clearly established even in the most straightforward case – that of simple action descriptions. As a first step in establishing these basic constraints, I sought to clarify the role of motor context on embodied perspective-taking in this most straightforward case. Whether the results obtained in Experiments 1 – 6 will generalise to more naturalistic language comprehension, with an online measure of processing and more complex linguistic stimuli, is not yet clear.

Certainly, evidence for a role of motor activation in abstract or figurative language is mixed. On the one hand, strongly embodied theories of language, in which motor resonance is required for comprehension, seek to show that abstract language involves the same motor circuits as concrete action descriptions (e.g. Glenberg et al., 2008), and there is some evidence from fMRI studies to support the view that motor activation occurs in response to metaphorical language (e.g., *the woman grasped the idea*; Desai et al., 2011), and idiomatic language (e.g., *the man kicked the bucket*, Boulenger, Hauk, & Pülvermüller, 2009). On the other hand, many studies aiming to demonstrate a role of the motor system in action-language understanding have used abstract or figurative language as controls, and found no evidence of motor activation in these control conditions (Aziz-Zedoh et al., 2006; Raposo et

al., 2009). A recent fMRI study found evidence for motor activation in the processing of metaphorical, but not idiomatic language (Cacciari et al., 2011).

It is not yet clear whether the view of motor resonance as a form of priming (section 2.3.5) predicts an effect in non-concrete language, but what is clear is that, under the priming account, such an effect is neither necessary nor sufficient for comprehension to occur. The results from Experiments 1 – 6 imply that motor resonance may only interact with language processing when an action is being planned concurrently during sentence processing. The role of motor resonance in naturalistic language comprehension may therefore depend on the extent to which a comprehender formulates a motor plan during naturalistic language use.

8.3.5. Motor resonance and prediction

This thesis did not explicitly test for a predictive role for motor resonance. However, the suggestion that the language- action interaction may involve prospective simulation, rather than retrospective simulation, adds weight to recent suggestions that this may be the case. To my knowledge, no published article has explicitly tested this hypothesis; doing so seems like a natural next step in the embodied cognition research programme. Throughout this thesis, I have argued that motor resonance is not necessary for language comprehension, but serves to prime one sort of interpretation over another when it is present. However, it is possible that motor resonance does more than prime participants to select a particular agent's perspective (e.g., left-handed, right-handed) in an otherwise ambiguous situation such as the forced-choice sentence-picture matching task. Motor resonance might also be useful to the comprehender, by allowing her to better predict upcoming speech. This interpretation is more in line with Taylor and Zwaan's notion of fault-tolerant comprehension than with a functionally neutral priming account: motor activation is not necessary to language comprehension, but when present, improves comprehension through better prediction. Recent work by Pickering and Garrod (2007, 2013) has begun to establish a framework for

understanding how comprehenders might use communicative motor resonance (i.e., activation in response to the articulations involved in producing speech) to predict what a speaker is going to say and thus aid their comprehension. Future work should aim to extend this framework by seeking to integrate referential motor resonance into such an account.

For example, an obvious question concerns the relationship between the two types of motor resonance. Research suggests that comprehenders automatically imitate the articulatory gestures of incoming speech (Fadiga et al., 2002); Pickering and Garrod (2007) therefore suggest that comprehenders predict upcoming speech using their own speech production systems (communicative motor resonance). However, research also suggests that language comprehension involves fast and apparently automatic prediction based on referential meaning, implying something akin to referential motor resonance (Altmann & Kamide, 1999; Kamide, 2003). How then, do these two levels of prediction interact with one another? One possibility is that the two sources of motor resonance (referential versus communicative) occur in different situations (isolated comprehension versus dialogue or production) and therefore do not interact. For example, I noted in section 8.3.5 above, that whether or not there is an effect of motor resonance on language may depend on whether or not the language user has formulated a particular motor plan during comprehension. In Experiments 1-9, I encouraged participants to form a motor plan to move a particular hand. However, in a more natural dialogue situation, comprehenders may produce motor plans of what their interlocutor is going to say, based on their own production system (Pickering & Garrod, 2013). Thus, the manual action plans formulated during sentence processing in Experiments 1-9 may translate to forward models of language production in situations where more than one interlocutor is present. This would involve a transition from referential motor resonance in isolated comprehension tasks such as used in Experiments 1-9 thesis (and throughout embodiment research) to communicative motor resonance in dialogue settings.

A second possibility is that both levels of motor resonance operate simultaneously. One question for future research is to establish the degree of computational difficulty that

simultaneous use of the two systems would entail, and whether one type of motor resonance precedes the other. Of course, language comprehension is not unique in requiring simultaneous levels of motor activation. Trivially, any time in which we conduct two actions at once, such as walking while eating an apple, requires us simultaneously predict and react to feedback from two different actions involving different effectors and goals. However, language poses particularly interesting issues, because of the fact that in only one of the levels (communicative motor resonance), is there any direct perceptual or proprioceptive feedback. The other level (referential motor resonance) operates without motor feedback (but see Pickering & Garrod, 2013, for an account of how referential motor resonance might involve cognitive feedback). The lack of motor feedback is of particular interest given its posited role in prediction of motor actions (Grush, 2004; Wolpert, Ghahramani, & Jordan, 1995). Investigating the possible predictive role of referential motor resonance, may therefore inform not only embodied cognition, but wider language processing and motor planning as a whole.

8.3.6. A functional role for motor resonance?

One of the aims of this thesis was to investigate the possibility of a functional role for referential motor resonance in action language understanding. The results from Experiments 1 – 6 suggest that planning to perform an action with a particular hand affects the way in which comprehenders interpret action sentences. However, these effects are found in a task requiring a manual response, and in which image selection occurs after sentence processing is complete. In Experiments 7 – 9, I used a task measuring online sentence processing in the absence of a manual response. Some findings (i.e., that participants are faster to perform spoken sensibility judgments when planning a particular type of action) could be interpreted as weak evidence for a general effect of motor activation on language processing. However, these findings could equally be interpreted as an effect of task difficulty on processing in general.

The only finding from Chapter 7 that does suggest some effect of action on language, is the fact that right-handed participants were more likely to make a correct congruency judgment when planning to execute an action with their dominant right hand, compared with their left hand. There was however, no effect of hand on the latency of congruency judgments. Taken together, the results from this thesis show that motor activation from current motor plan affects the way people interpret action sentences, in offline tasks, and that this effect interacts with sentence pronoun so that comprehenders are more likely to adopt an embodied agent's perspective on first-person versus third-person sentences. However, evidence for an effect of action on language in online comprehension is much weaker, with accuracy rates affected only when the task requires an explicit comparison between the action described in the sentence, and the action the comprehender is planning to make. In tasks where this comparison is irrelevant (Experiments 7 and 8), there is no effect of hand on accuracy rates. Based on the current data therefore, there is insufficient evidence to answer the question: to what extent do comprehenders routinely adopt an embodied perspective on action language, and to what extent do they adopt a non-embodied perspective?

In section 3.6, I argued that it is important for embodied approaches to language to allow for a non-embodied perspective, otherwise they must make the problematic claim that motor resonance is necessary for comprehension. But under what circumstances would we adopt an embodied perspective, and under what circumstances would we adopt a non-embodied perspective? The results from Experiments 1 - 6 suggest that, in the forced-choice task at least, comprehenders adopted an embodied perspective (namely, that of the agent) in line with their current motor plan. If, during sentence processing, there was no motor plan to move a particular hand, then an unembodied perspective was adopted (i.e., participants did not show a preference for the image congruent with their response hand, which would indicate an embodied agent's perspective; and they did not show a preference for right-handed images, which would indicate an embodied observer's perspective, in line with their

visual experience). Once a motor plan is formed, the Spatial Grounding Hypothesis predicts that the effect of this plan will be stronger following self-referential language (which provides a means of grounding the simulated action in space, since the agent of the action is interpreted as the comprehender) compared with third-person language. Results from Experiments 1 - 6 mostly support this claim, with the preference for images that matched participant's current motor plan interacting with pronoun in Experiments 2a, 2b, 5, and 6 (see section 8.3.2 for discussion).

Throughout the thesis, I have argued that motor resonance is not necessary for comprehension. The results of Experiments 1 – 9 support this claim by finding an effect of motor activation on sentence interpretation *only in certain circumstances* (i.e., when an action plan was present during sentence processing). Thus, motor activation is not necessary for language comprehension, unless we change our definition of what it means to comprehend a sentence. I have argued that conceiving of motor activation as a type of priming allows us to acknowledge the fact that motor resonance can influence comprehension, without being either necessary or sufficient for comprehension to occur. In this simple priming account, motor activation primes the comprehender to adopt a particular perspective, but does not serve any useful function. The idea of fault-tolerant comprehension on the other hand (Taylor & Zwaan, 2009, 2013), proposes that although motor activation is not necessary for language comprehension, it does provide a useful function: comprehension is more efficient in the presence of motor resonance than in its absence. I suggested in section 8.3.5 above that one conduit for this increased efficiency might be improved prediction. It is not clear whether fault-tolerant comprehension applies to abstract and metaphorical language as well as concrete action descriptions: one possibility is that comprehension of abstract language is generally impoverished compared with concrete language for which simulations are readily available, thus explaining the well-documented differences in processing for concrete versus abstract words (see Paivio, 1991 for a review).

The idea that motor resonance allows better prediction might explain the finding that participants were more likely to adopt an embodied agent's perspective for self-referential language compared with third-person language. One of the reasons for predicting a difference between first- and third-person language was the evidence that first-person (self-generated) actions are more tightly connected to the motor system than third-person (other generated) actions (see section 2.2.5). Specifically, people are better at predicting their own actions than those of other people (Knoblich & Flach, 2001). Therefore, if motor resonance plays a predictive role, and if first-person actions result in better predictions, then it follows that there might be a stronger effect of motor resonance in first-person than in third-person language. This prediction is borne out by the findings from Experiments 2a -6. However, if motor resonance does aid prediction (presumably in an online capacity, since retrospective prediction would appear to be of little use), is not clear why the effects of action plan were limited to the forced-choice task in Experiments 2a – 6. If language comprehension is fault-tolerant in the sense that comprehension is achievable, but less optimal without the predictive advantage supplied by referential motor resonance, then we might expect to see some effect of motor resonance in the sensibility judgement task (Experiments 7 – 9). No such effect was obtained.

One explanation for this pattern of results is that, as Zwaan and Taylor (2006) point out, sensibility judgements are made once participants have finished reading the sentence, and might well consist of post-sentential re-imaginings, in order to complete the task. Any predictive advantage afforded by referential motor resonance may occur earlier in the sentence – perhaps captured by the localised, verb-specific effects found by Zwaan & Taylor (2006; Taylor & Zwaan, 2008). Eye-tracking methods should therefore be used to test this possibility using an online measure of sentence processing. However, the argument that the null effects in Experiments 7 – 9 were due to us tapping into a later, post-sentential reimagining rather than true online motor resonance, can only provide a partial explanation,

since the task in Experiments 1 – 6 also involved a post-sentential judgement, but did pick up an effect of response hand (see section 7.7.1 and 7.7.2 for further discussion of the null effects in Experiments 7-9).

In this thesis I have argued against motor resonance (and therefore, an embodied perspective) being necessary for action language understanding; I have also provided some evidence that, using an off-line task, an embodied perspective is adopted in line with the following constraints: (1) the comprehender has a motor plan present during sentence comprehension; (2) the comprehender is able to sufficiently ground the action simulation in space, for example by reference to her own body. I have found little evidence to suggest that this embodied perspective plays a functional role in action language comprehension.

8.4. Concluding remarks

In this thesis, I set out to specify some of the constraints governing the interaction between language and action. Through a series of sentence-picture matching studies, I have demonstrated that comprehenders tend to adopt an embodied agent's perspective on action sentences, that the evidence for this perspective is stronger following self-referential language than third-person language, and that this perspective appears to be grounded in the comprehender's current motor plan rather than their motor experience. Contrary to the predictions of the Body-Specificity Hypothesis, I found no effect of dominant hand (i.e., motor experience) in the absence of manual responses. However, in a different task, which explicitly required participants to judge the congruency between their planned action and the sentence they were reading, there was some evidence of an effect of dominant hand on sentence processing. In other words, comprehenders seem to assign a handedness representation to action language only when required to do so by the task. These results support weak embodied accounts in which the way we interpret action sentences is influenced by activation of the motor system, and suggest that the interaction between action

and language is sensitive to the distinction between self and other. The results provide no evidence in favour of strong embodied accounts, where motor activation is necessary for comprehension to occur.

REFERENCES

- Abell, F., Happé, F., & Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cognitive Development, 15*, 1–16.
- Adams, F. (2010). Embodied cognition. *Phenomenology and the Cognitive Sciences, 9*, 619–628.
- Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience, 11*, 1109–1116.
- Aglioti, S. M., & Pazzaglia, M. (2010). Representing actions through their sound. *Experimental Brain Research, 206*, 141–151.
- Alaerts, K., Heremans, E., Swinnen, S. P., & Wenderoth, N. (2009). How are observed actions mapped to the observer's motor system? Influence of posture and perspective. *Neuropsychologia, 47*, 415–22.
- Alaerts, K., Swinnen, S. P., & Wenderoth, N. (2009). Is the human primary motor cortex activated by muscular or direction-dependent features of observed movements? *Cortex, 45*, 1148–1155.
- Altmann, G., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*, 247–264.
- Altschuler, E. L., Vankok, A., Hubbard, E. M., Roberts, E., Ramachandran, V. S., & Pineda, J. A. (2000). Mu wave blocking by observation of movement and its possible use as a tool to study theory of other minds. In *Society of Neuroscience* (p. 68).
- Amazeen, E. L., Amazeen, P. G., Treffner, P. J., & Turvey, M. T. (1997). Attention and handedness in bimanual coordination dynamics. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 1552–1560.
- Amazeen, E. L., Ringenbach, S. D., & Amazeen, P. G. (2005). The effects of attention and handedness on coordination dynamics in a bimanual Fitts' law task. *Experimental brain research, 164*, 484–499.
- Anderson, M. L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences, 33*, 245–66; discussion 266–313.
- Anderson, S. J., Yamagishi, N., & Karavia, V. (2002). Attentional processes link perception and action. *Proceedings of the Royal Society of London, B, 269*, 1225–1232.
- Anquetil, T., & Jeannerod, M. (2007). Simulated actions in the first and in the third person perspectives share common representations. *Brain Research, 1130*, 125–129.

- Aravena, P., Hurtado, E., Riveros, R., Cardona, J. F., Manes, F., & Ibáñez, A. (2010). Applauding with closed hands: Neural signature of action-sentence compatibility effects. *PLOS ONE*, 5, e11751.
- Arbib, M. A. (2005). From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, 28(2), 105–124; discussion 125–167.
- Arévalo, A., Baldo, J. V., & Dronkers, N. F. (2012). What do brain lesions tell us about theories of embodied semantics and the human mirror neuron system? *Cortex*, 48, 242–254.
- Ashby, F., Gregory, B., Turner, O., & Horvitz, J. C. (2010). Cortical and basal ganglia contributions to habit learning and automaticity. *Trends in Cognitive Sciences* 14, 208–215.
- Aziz-Zadeh, L., & Damasio, A. (2008). Embodied semantics for actions: Findings from functional brain imaging. *Journal of Physiology*, 102, 35–39.
- Aziz-Zadeh, L., Maeda, F., Zaidel, E., Mazziotta, J., & Iacoboni, M. (2002). Lateralization in motor facilitation during action observation: A TMS study. *Experimental Brain Research*, 144, 127–131.
- Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G., & Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology*, 16, 1818–1823.
- Baayen, R. H. (2008). *Analysing linguistic data*. Cambridge: Cambridge University Press.
- Bak, T. H., O'Donovan, D. G., Xuereb, J. H., Boniface, S., & Hodges, J. R. (2001). Selective impairment of verb processing associated with pathological changes in Brodmann areas 44 and 45 in the motor neurone disease-dementia-aphasia syndrome. *Brain*, 124, 103–120.
- Barclay, J. R. (1973). The role of comprehension in remembering. *Cognitive Psychology*, 254, 229–254.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577–609.
- Barsalou, L. W. (2003). Situated Simulation in the Human Conceptual System. *Language and Cognitive Processes*, 18, 513–562.
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617–645.

- Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, 364, 1281–1289.
- Barsalou, L. W., Kyle Simmons, W., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, 7, 84–91.
- Bates, D. M., & Maechler, M. (2009). *lme4: Linear mixed-effects models using Eigen and R classes*. R package version 0.999375-32.
- Bavelas, J., Black, A., Lemery, C., & Mullet, J. (1986). “I show how you feel”: Motor mimicry as a communicative act. *Journal of Personality and Social Psychology*, 50, 322–329.
- Beardsworth, T., & Buckner, T. (1981). The ability to recognise oneself from a video recording of one’s movements without seeing one’s body. *Bulletin of the Psychonomic Society*, 18, 19–22.
- Beilock, S. L., Lyons, I. M., Mattarella-Micke, A., Nusbaum, H. C., & Small, S. L. (2008). Sports experience changes the neural processing of action language. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 13269–13273.
- Bekkering, H., & Neggers, S. F. W. (2002). Visual search is modulated by action intentions. *Psychological Science*, 13, 370–374.
- Bergen, B., & Wheeler, K. (2010). Grammatical aspect and mental simulation. *Brain and Language*, 112, 150–158.
- Bernieri, F. J. (1988). Coordinated movement and rapport in teacher-student interactions. *Journal of Nonverbal Behavior*, 12, 120–138.
- Beveridge, M. E. L., & Pickering, M. J. (2013). Perspective taking in language: integrating the spatial and action domains. *Frontiers in Human Neuroscience*, 7, 577.
- Bidet-Ildei, C., Chauvin, A., & Coello, Y. (2010). Observing or producing a motor action improves later perception of biological motion: evidence for a gender effect. *Acta Psychologica*, 134, 215–224.
- Bidet-Ildei, C., Sparrow, L., & Coello, Y. (2011). Reading action word affects the visual perception of biological motion. *Acta Psychologica*, 137, 330–334.
- Binkofski, F., Buccino, G., Posse, S., Seitz, R. J., Rizzolatti, G., & Freund, H. (1999). A fronto-parietal circuit for object manipulation in man: evidence from an fMRI-study. *The European Journal of Neuroscience*, 11, 3276–3286.
- Bird, G., Leighton, J., Press, C., & Heyes, C. (2007). Intact automatic imitation of human and robot actions in autism spectrum disorders. *Proceedings of the Royal Society of London, B*, 274, 3027–3031.

- Black, M., Nickels, L., & Byng, S. (1991). Patterns of sentence processing deficit: Processing simple sentences can be a complex matter. *Journal of Neurolinguistics*, 6, 79–101.
- Blakemore, S.-J., & Frith, C. (2005). The role of motor contagion in the prediction of action. *Neuropsychologia*, 43, 260–267.
- Borghi, A. M. (2004). Object concepts and action: Extracting affordances from objects parts. *Acta Psychologica*, 115, 69–96.
- Borghi, A. M. (2013). Language comprehension: Action, affordances and goals. In Y. Coello & A. Bartolo (Eds.), *Language and action in cognitive neuroscience* (pp. 125–144). Hove: Psychology Press.
- Borghi, A. M., Glenberg, A. M., & Kaschak, M. P. (2004). Putting words in perspective. *Memory and Cognition*, 32, 863–873.
- Borghi, A. M., & Riggio, L. (2009). Sentence comprehension and simulation of object temporary, canonical and stable affordances. *Brain Research*, 1253, 117–128.
- Borghi, A. M., & Scorolli, C. (2009). Language comprehension and dominant hand motion simulation. *Human Movement Science*, 28, 12–27.
- Borreggine, K. L., & Kaschak, M. P. (2006). The action-sentence compatibility effect: It's all in the timing. *Cognitive Science*, 30, 1097–1112.
- Boulenger, V., Hauk, O., & Pulvermüller, F. (2009). Grasping Ideas with the Motor System: Semantic Somatotopy in Idiom Comprehension. *Cerebral Cortex*, 19, 1905–1914.
- Boulenger, V., Mechtouff, L., Thobois, S., Broussolle, E., Jeannerod, M., & Nazir, T. a. (2008). Word processing in Parkinson's disease is impaired for action verbs but not for concrete nouns. *Neuropsychologia*, 46, 743–756.
- Boulenger, V., Roy, A., Paulignan, Y., Deprez, V., Jeannerod, M., & Nazir, T. (2006). Cross-talk between language processes and overt motor behavior in the first 200 msec of processing. *Journal of Cognitive Neuroscience*, 18, 1607–1615.
- Boulenger, V., Silber, B. Y., Roy, A. C., Paulignan, Y., Jeannerod, M., & Nazir, T. a. (2008). Subliminal display of action words interferes with motor planning: a combined EEG and kinematic study. *Journal of Physiology*, 102, 130–136.
- Bouquet, C. A., Shipley, T. F., Capa, R. L., & Marshall, P. J. (2011). Motor contagion: goal-directed actions are more contagious than non-goal-directed actions. *Experimental Psychology*, 58, 71–78.
- Branigan, H. P., Pickering, M. J., & McLean, J. F. (2005). Priming prepositional-phrase attachment during comprehension. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 31, 468–481.
- Bransford, J. D., Barclay, J. R., & Franks, J. (1972). Sentence Memory: A constructive versus interpretive approach. *Cognitive Psychology*, 3, 193–209.

- Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica*, *106*, 3–22.
- Brass, M., Bekkering, H., Wohlschläger, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: Comparing symbolic, spatial, and imitative cues. *Brain and Cognition*, *44*, 124–143.
- Brunyé, T. T., Ditman, T., Mahoney, C. R., Augustyn, J. S., & Taylor, H. A. (2009). When You and I Share Perspectives: Pronouns modulate perspective taking during narrative comprehension. *Psychological Science*, *20*, 27–32.
- Brunyé, T. T., Ditman, T., Mahoney, C. R., & Taylor, H. A. (2011). Better you than I: Perspectives and emotion simulation during narrative comprehension. *Journal of Cognitive Psychology*, *23*, 659–666.
- Brunyé, T. T., Gardony, A., Mahoney, C. R., & Taylor, H. A. (2012). Body-specific representations of spatial location. *Cognition*, *123*, 229–339.
- Bryden, M. P. (1977). Measuring handedness with questionnaires. *Neuropsychologia*, *15*, 617–624.
- Bub, D. N., Masson, M. E. J., & Cree, G. S. (2008). Evocation of functional and volumetric gestural knowledge by objects and words. *Cognition*, *106*, 27–58.
- Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., ... Parma, I.-. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience*, *13*, 400–405.
- Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., & Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: A combined TMS and behavioral study. *Cognitive Brain Research*, *24*, 355–363.
- Buccino, Giovanni, Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., ... Rizzolatti, G. (2004). Neural circuits involved in the recognition of actions performed by nonconspecifics: an FMRI study. *Journal of Cognitive Neuroscience*, *16*, 114–126.
- Buxbaum, L. J., Kyle, K. M., Tang, K., & Detre, J. A. (2006). Neural substrates of knowledge of hand postures for object grasping and functional object use: evidence from fMRI. *Brain Research*, *1117*, 175–185.
- Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: An FMRI study with expert dancers. *Cerebral Cortex*, *15*, 1243–1249.
- Calvo-Merino, B., Ehrenberg, S., Leung, D., & Haggard, P. (2010). Experts see it all: configural effects in action observation. *Psychological Research*, *74*, 400–406.
- Calvo-Merino, B., Grèzes, J., Glaser, D. E., Passingham, R. E., & Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Current Biology*, *16*, 1905–1910.

- Campanella, F., Sandini, G., & Morrone, M. C. (2011). Visual information gleaned by observing grasping movement in allocentric and egocentric perspectives. *Proceedings of the Royal Society of London, B*, 278, 2142–2149.
- Caplan, D., & Waters, G. (1997). Determinants of sentence comprehension in aphasic patients in sentence-picture matching tasks. *Journal of Speech, Language, and Hearing Research*, 40, 542–555.
- Casasanto, D. (2011). Different bodies, different minds: The body specificity of language and thought. *Current Directions in Psychological Science*, 20, 378–383.
- Casasanto, D. (2009). Embodiment of abstract concepts: good and bad in right- and left-handers. *Journal of Experimental Psychology: General*, 138, 351–67.
- Casasanto, D., & Chrysikou, E. G. (2011). When left is “right”. Motor fluency shapes abstract concepts. *Psychological Science*, 22, 419–422.
- Casasanto, D., & Henetz, T. (2012). Handedness shapes children’s abstract concepts. *Cognitive Science*, 36, 359–372.
- Casile, A., & Giese, M. A. (2006). Nonvisual motor training influences biological motion perception. *Current Biology*, 16, 69–74.
- Castiello, U., Lusher, D., Morena, M., Edwards, M., & Humphreys, G. (2002). Observing a human or robotic hand grasping an object: Differential motor priming effect. In Wolfgang Prinz & B. Hommel (Eds.), *Common mechanisms in perception and action. Attention and performance XIX*. (pp. 315–333). Oxford: Oxford University Press.
- Catmur, C., Walsh, V., & Heyes, C. (2007). Sensorimotor learning configures the human mirror system. *Current Biology*, 17, 1527–1531.
- Catmur, C., Walsh, V., & Heyes, C. (2009). Associative sequence learning: the role of experience in the development of imitation and the mirror system. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 364, 2369–2380.
- Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behaviour link and social interaction. *Journal of Personality and Social Psychology*, 76, 893–910.
- Chase, W. G., & Clark, H. H. (1971). Semantics in the perception of verticality. *British Journal of Psychology*, 62, 311–326.
- Chase, W. G., & Clark, H. H. (1972). Mental operations in the comparison of sentences and pictures. In L. W. Gregg (Ed.), *Cognition in learning and memory* (pp. 205–252). London: Wiley.
- Chatterjee, A., Southwood, M. H., & Basilico, D. (1999). Verbs, events and spatial representations. *Neuropsychologia*, 37, 395–402.
- Chatterjee, A. (2010). Disembodying cognition. *Language and Cognition*, 2, 79–116.

- Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology*, 42, 368–407.
- Clark, A. (1999). An embodied cognitive science? *Trends in Cognitive Sciences*, 3, 345–351.
- Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.
- Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology*, 517, 472–517.
- Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22, 1–39.
- Cochin, S., Barthelemy, C., Lejeune, B., Roux, S., & Martineau, J. (1998). Perception of motion and qEEG activity in human adults. *Electroencephalography and Neurophysiology*, 107, 287–295.
- Cochin, S., Barthelemy, C., & Roux, S. (1999). Observation and execution of movement : Similarities demonstrated by quantified electroencephalography. *European Journal of Cognitive Psychology*, 11, 1839–1842.
- Cohen, G. (1969). Pattern recognition: Differences between matches patterns to patterns and matching descriptions to patterns. *Journal of Experimental Psychology*, 82, 427–434.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum
- Corballis, M. C. (2013). Towards a Darwinian perspective on language evolution. In Y. Coello & A. Bartolo (Eds.), *Language and action in cognitive neuroscience* (pp. 33–58). Hove: Psychology Press.
- Cotelli, M., Borroni, B., Manenti, R., Alberici, A., Calabria, M., Agosti, C., ... Cappa, S. F. (2006). Action and object naming in frontotemporal dementia, progressive supranuclear palsy, and corticobasal degeneration. *Neuropsychology*, 20, 558–565.
- Cotelli, M., Borroni, B., Manenti, R., Zanetti, M., Arévalo, A., Cappa, S. F., & Padovani, A. (2007). Action and object naming in Parkinson's disease without dementia. *European Journal of Neurology*, 14, 632–637.
- Craighero, L., Bello, A., Fadiga, L., & Rizzolatti, G. (2002). Hand action preparation influences the responses to hand pictures. *Neuropsychologia*, 40, 492–502.
- Craighero, L., Fadiga, L., Rizzolatti, G., & Umiltà, M. A. (1999). Action for perception: A motor-visual attentional effect. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1673–1692.

- Craigthero, L., Fadiga, L., Umiltà, C. A., & Rizzolatti, G. (1996). Evidence for visuomotor priming effect. *NeuroReport*, 8, 347–349.
- Creem, S. H., & Proffitt, D. R. (2001). Grasping objects by their handles: A necessary interaction between cognition and action. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 218–228.
- Crescentini, C., Mengotti, P., Grecucci, A., & Rumiati, R. I. (2011). The effect of observed biological and non biological movements on action imitation: an fMRI study. *Brain Research*, 1420, 80–92.
- Csibra, G. (2008). Action mirroring and action understanding: An alternative account. In P. Haggard, Y. Rossetti, & M. Kawato (Eds.), *Sensorimotor foundations of higher cognition: Attention and performance XXII* (pp. 435–459). Oxford: Oxford University Press.
- Cutting, J., & Kozlowski, L. (1977). Recognising friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, 9, 353–356.
- D'Ausilio, A., Craigthero, L., & Fadiga, L. (2012). The contribution of the frontal lobe to the perception of speech. *Journal of Neurolinguistics*, 25, 328–335.
- Dalla Volta, R., Gianelli, C., Campione, G. C., & Gentilucci, M. (2009). Action word understanding and overt motor behavior. *Experimental Brain Research*, 196, 403–412.
- Daprati, E., Franck, N., Georgieff, N., Proust, J., Pacherie, E., Dalery, J., & Jeannerod, M. (1997). Looking for the agent: an investigation into consciousness of action and self-consciousness in schizophrenic patients. *Cognition*, 65, 71–86.
- Daprati, E., Wriessnegger, S., & Lacquaniti, F. (2007a). Knowledge of one's kinematics improves perceptual discrimination. *Consciousness and Cognition*, 16, 178–188.
- Daprati, E., Wriessnegger, S., & Lacquaniti, F. (2007b). Kinematic cues and recognition of self-generated actions. *Experimental Brain Research*, 177, 31–44.
- Dapretto, M., & Bookheimer, S. Y. (1999). Form and content: dissociating syntax and semantics in sentence comprehension. *Neuron*, 24, 427–432.
- Davidson, R. J. (1992). Anterior cerebral asymmetry and the nature of emotion. *Brain and Cognition*, 20, 125–151.
- De Jong, R. (1993). Multiple bottlenecks in overlapping task performance. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 965–980.
- De Jong, R. (1995). The Role of Preparation in Overlapping-task Performance. *The Quarterly Journal of Experimental Psychology*, 48, 2–25.
- De Jong, R., & Sweet, J. B. (1994). Preparatory strategies in overlapping task performance. *Attention, Perception and Psychophysics*, 55, 142–151.

- De Vega, M., Moreno, V., & Castillo, D. (2013). The comprehension of action-related sentences may cause interference rather than facilitation on matching actions. *Psychological Research*, 77, 20–30.
- De Vignemont, F., & Haggard, P. (2008). Action observation and execution: What is shared? *Social Neuroscience*, 3, 421–433.
- Decety, J., & Grèzes, J. (1999). Neural mechanisms subserving the perception of human actions. *Trends in Cognitive Sciences*, 3, 172–178.
- Decety, J. (2002). Neurophysiological evidence for simulation of action. In J. Dokic & J. Proust (Eds.), *Simulation and knowledge of action* (pp. 53–72). Amsterdam: John Benjamins.
- Decety, J., & Chaminade, T. (2003). When the self represents the other: A new cognitive neuroscience view on psychological identification. *Consciousness and Cognition*, 12, 577–596.
- Decety, J., Perani, D., Jeannerod, M., Bettinardi, V., Tadary, T., Woods, R., ... Fazio, F. (1994). Mapping motor representations with positron emission tomography. *Nature*, 371, 600–602.
- Decety, J., & Sommerville, J. A. (2003). Shared representations between self and other: a social cognitive neuroscience view. *Trends in Cognitive Sciences*, 7, 527–533.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93, 283–321.
- Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, 176–180.
- Diefenbach, C., Rieger, M., Massen, C., & Prinz, W. (2013). Action-sentence compatibility: The role of action effects and timing. *Frontiers in Psychology*, 4, e272.
- Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science*, 11, 86–89.
- Dimberg, U., & Thunberg, M. (1998). Rapid facial reactions to emotional facial expressions. *Scandinavian Journal of Psychology*, 39, 39–45.
- Dinstein, I., Hasson, U., Rubin, N., & Heeger, D. J. (2007). Brain areas selective for both observed and executed movements. *Journal of Neurophysiology*, 98, 1415–1427.
- Dinstein, I., Thomas, C., Behrmann, M., & Haeger, D. (2008). A mirror up to nature. *Current Biology*, 18, 3–8.
- Ditman, T., Brunyé, T. T., Mahoney, C. R., & Taylor, H. A. (2010). Simulating an enactment effect: Pronouns guide action simulation during narrative comprehension. *Cognition*, 115, 172–178.

- Dayan, E., & Cohen, L. G. (2011). Neuroplasticity subserving motor skill learning. *Neuron*, 72, 443–454.
- Doyon, J., & Benali, H. (2005). Reorganization and plasticity in the adult brain during learning of motor skills. *Current Opinion in Neurobiology*, 15, 161–167.
- Dronkers, N. F., Plaisant, O., Iba-Zizen, M. T., & Cabanis, E. A. (2007). Paul Broca's historic cases: High resolution MR imaging of the brains of Leborgne and Lelong. *Brain*, 130, 1432–1441.
- Drost, U. C., Rieger, M., & Prinz, W. (2007). Instrument specificity in experienced musicians. *Quarterly Journal of Experimental Psychology*, 60, 527–533.
- Duran, N. D., Dale, R., & Kreuz, R. J. (2011). Listeners invest in an assumed other's perspective despite cognitive cost. *Cognition*, 121, 22–40.
- Ellis, R., & Tucker, M. (2000). Micro-affordance : The potentiation of components of action by seen objects. *British Journal of Psychology*, 91, 451–471.
- Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles : a TMS study. *European Journal of Neuroscience*, 15, 399–402.
- Fadiga, L., Craighero, L., & D'Ausilio, A. (2009). Broca's area in language, action, and music. *Annals of the New York Academy of Sciences*, 1169, 448–458.
- Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology*, 73, 2608–2611.
- Fagioli, S., Ferlazzo, F., & Hommel, B. (2007). Controlling attention through action: observing actions primes action-related stimulus dimensions. *Neuropsychologia*, 45, 3351–3355.
- Fagioli, S., Hommel, B., & Schubotz, R. I. (2007). Intentional control of attention: action planning primes action-related stimulus dimensions. *Psychological Research*, 71, 22–29.
- Fernandino, L., & Iacoboni, M. (2010). Are cortical motor maps based on body parts or coordinated actions? Implications for embodied semantics. *Brain and Language*, 112, 44–53.
- Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, 47, 164–203.
- Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, 11, 11–15.
- Ferreira, F., & Patson, N. D. (2007). The “good enough” approach to language comprehension. *Language and Linguistics Compass*, 1, 71–83.

- Fischer, M. H., & Zwaan, R. A. (2008). Embodied language: A review of the role of the motor system in language comprehension. *Quarterly Journal of Experimental Psychology*, 61, 825–850.
- Flach, R., Knoblich, G., & Prinz, W. (2004). Recognizing one's own clapping: The role of temporal cues. *Psychological Research*, 69, 147–156.
- Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the Level 1-Level 2 distinction. *Developmental Psychology*, 17, 99–103.
- Fodor, J. (1975). *The language of thought*. Harvard: Harvard University Press.
- Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.
- Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: From action organization to intention understanding. *Science*, 308, 662–667.
- Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and Performance XII: The Psychology of Reading* (pp. 601–681). Hillsdale, NJ: Erlbaum.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, 6, 78–84.
- Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, 13, 361–77.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 11, 593–609.
- Gallese, V. (2008). Mirror neurons and the social nature of language: the neural exploitation hypothesis. *Social Neuroscience*, 3, 317–33.
- Gallese, V. (2009). Motor abstraction: A neuroscientific account of how action goals and intentions are mapped and understood. *Psychological Research*, 73, 486–498.
- Gallese, V., & Goldman, A. (1998). Mirror neurons and the mind-reading. *Trends in Cognitive Sciences*, 2, 493–501.
- Gallese, V., & Lakoff, G. (2005). The Brain's concepts: the role of the Sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, 22, 455–479.
- Gambi, C., & Pickering, M. J. (2011). A cognitive architecture for the coordination of utterances. *Frontiers in Psychology*, 2, 1–14.
- Gardner, M. R., & Potts, R. (2010). Hand dominance influences the processing of observed bodies. *Brain and Cognition*, 73, 35–40.

- Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press.
- Gentilucci, M., Benuzzi, F., Bertolani, L., Daprati, E., & Gangitano, M. (2000). Language and motor control. *Experimental Brain Research*, 133, 468–490.
- Gentilucci, M., & Gangitano, M. (1998). Influence of automatic word reading on motor control. *European Journal of Neuroscience*, 10, 752–756.
- Gerardin, E., Sirigu, A., Lehericy, S., Poline, J. B., Gaymard, B., Marsault, C., ... Le Bihan, D. (2000). Partially overlapping neural networks for real and imagined hand movements. *Cerebral Cortex*, 10, 1093–1104.
- Gianelli, C., Farnè, A., Salemme, R., Jeannerod, M., & Roy, A. C. (2011). The agent is right: when motor embodied cognition is space-dependent. *PLOS ONE*, 6(9), e25036.
- Gianelli, C., Scorolli, C., & Borghi, A. M. (2013). Acting in perspective: The role of body and language as social tools. *Psychological Research*, 77, 40–52.
- Gibbs, R. W. (2006). *Embodiment and cognitive science*. Cambridge: Cambridge University Press.
- Gibson, J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing: Toward an ecological perspective* (pp. 67–82). Hillsdale, NJ: Lawrence Erlbaum.
- Gilead, M., Liberman, N., & Maril, A. (2013). The language of future-thought: an fMRI study of embodiment and tense processing. *NeuroImage*, 65, 267–279.
- Gillmeister, H., Catmur, C., Liepelt, R., Brass, M., & Heyes, C. (2008). Experience-based priming of body parts: a study of action imitation. *Brain Research*, 1217, 157–170.
- Glenberg, A. M. (2007). Language and action: Creating sensible combinations of ideas. In G. Gaskell (Ed.), *The Oxford Handbook of Psycholinguistics* (pp. 361–370). Oxford: Oxford University Press.
- Glenberg, A. M., & Gallese, V. (2012). Action-based language: A theory of language acquisition, comprehension, and production. *Cortex*, 48, 905–922.
- Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, 9, 558–565.
- Glenberg, A. M., & Robertson, D. (2000). Symbol Grounding and Meaning: A Comparison of High-Dimensional and Embodied Theories of Meaning. *Journal of Memory and Language*, 43, 379–401.
- Glenberg, A. M., Sato, M., & Cattaneo, L. (2008). Use-induced motor plasticity affects the processing of abstract and concrete language. *Current Biology*, 18, 290–291.

- Glenberg, A. M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D., & Buccino, G. (2008). Processing abstract language modulates motor system activity. *Quarterly Journal of Experimental Psychology*, *61*, 905–919.
- Glover, S., & Dixon, P. (2002). Semantics affect the planning but not control of grasping. *Experimental Brain Research*, *146*, 383–387.
- Glover, S., Rosenbaum, D. A., Graham, J., & Dixon, P. (2004). Grasping the meaning of words. *Experimental Brain Research*, *154*, 103–108.
- Grafton, S. T., Arbib, M. A., Fadiga, L., & Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography. *Experimental Brain Research*, *112*, 103–111.
- Grèzes, J., Armony, J. L., Rowe, J., & Passingham, R. E. (2003). Activations related to “mirror” and “canonical” neurones in the human brain: an fMRI study. *NeuroImage*, *18*, 928–937.
- Grèzes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Human Brain Mapping*, *12*, 1–19.
- Grèzes, J., Frith, C. D., & Passingham, R. E. (2004). Inferring false beliefs from the actions of oneself and others : an fMRI study. *NeuroImage*, *21*, 744–750.
- Grush, R. (2004). The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and Brain Sciences*, *27*, 377–396; discussion 396–442.
- Hamilton, A., Wolpert, D., & Frith, U. (2004). Your own action influences how you perceive another person’s action. *Current Biology*, *14*, 493–498.
- Hari, R., & Salmelin, R. (1997). Human cortical oscillations : A neuromagnetic view through the skull. *Trends in Neurosciences*, *20*, 44–49.
- Haruno, M., Wolpert, D. M., & Kawato, M. (2001). Mosaic model for sensorimotor learning and control. *Neural computation*, *13*, 2201–2220.
- Haueisen, J., & Knösche, T. R. (2001). Involuntary motor activity in pianists evoked by music perception. *Journal of Cognitive Neuroscience*, *13*, 786–792.
- Hauk, O., Shtyrov, Y., & Pulvermüller, F. (2008). The time course of action and action-word comprehension in the human brain as revealed by neurophysiology. *Journal of physiology, Paris*, *102*, 50–58.
- Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*, 301–730.
- Hauk, O., & Pulvermüller, F. (2004). Neurophysiological distinction of action words in the fronto-central cortex. *Human brain mapping*, *21*, 191–201.

- Hauk, O., & Pulvermüller, F. (2011). The lateralization of motor cortex activation to action-words. *Frontiers in Human Neuroscience*, 5, 149.
- Hecht, H., Vogt, S., & Prinz, W. (2001). Motor learning enhances perceptual judgment: A case for action-perception transfer. *Psychological Research*, 65, 3–14.
- Helbig, H. B., Graf, M., & Kiefer, M. (2006). The role of action representations in visual object recognition. *Experimental Brain Research*, 174, 221–228.
- Heyes, C. (2010). Where do mirror neurons come from? *Neuroscience and Biobehavioral Reviews*, 34, 575–583.
- Heyes, C. (2011). Automatic imitation. *Psychological Bulletin*, 137, 463–483.
- Heyes, C., Bird, G., Johnson, H., & Haggard, P. (2005). Experience modulates automatic imitation. *Brain Research*, 22, 233–240.
- Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*, 21, 1229–1243.
- Hickok, G. (2010). The role of mirror neurons in speech perception and action word semantics. *Language and Cognitive Processes*, 25, 749–776.
- Hikosaka, O., Nakahara, H., Rand, M. K., Sakai, K., Lu, X., Nakamura, K., ... & Doya, K. (1999). Parallel neural networks for learning sequential procedures. *Trends in Neurosciences*, 22, 464–471.
- Hoedemaker, R. S., & Gordon, P. C. (2013). Embodied Language Comprehension: Encoding-Based and Goal-Driven Processes. *Journal of Experimental Psychology. General*.
- Hommel, B. (2011). The Simon effect as tool and heuristic. *Acta Psychologica*, 136, 189–202.
- Hommel, B., & Müsseler, J. (2006). Action-feature integration blinds to feature-overlapping perceptual events: Evidence from manual and vocal actions. *Quarterly Journal of Experimental Psychology*, 59, 509–523.
- Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, 24, 849–878.
- Honeck, R. P. (1973). Interpretive versus structural effects on semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 12, 448–455.
- Houk, J. C., & Wise, S. P. (1995). Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: their role in planning and controlling action. *Cerebral Cortex*, 5, 95–110.

- Huetting, F., & Altmann, G. (2007). Visual-shape competition during language-mediated attention is based on lexical input and not modulated by contextual appropriateness. *Visual Cognition*, 15, 985–1018.
- Huetting, F., & McQueen, J. M. (2011). The nature of the visual environment induces implicit biases during language-mediated visual search. *Memory & Cognition*, 39, 1068–1084.
- Hurley, S. (2001). Perception and action : Alternative views. *Synthese*, 129, 3–40.
- Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLOS ONE*, 3, e79.
- Intons-Peterson, M. J. (1983). Imagery paradigms: How vulnerable are they to experimenters' expectations?. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 394-412.
- Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience*, 9, 304-313.
- Jackson, P. L., Meltzoff, A. N., & Decety, J. (2006). Neural circuits involved in imitation and perspective-taking. *NeuroImage*, 31, 429–439.
- Jacob, P. (2008). What do mirror neurons contribute to human social cognition? *Mind and Language*, 23, 190–223.
- Jacob, P. (2009). The tuning-fork model of human social cognition: A critique. *Consciousness and Cognition*, 18, 229–243.
- Jacob, P., & Jeannerod, M. (2005). The motor theory of social cognition: a critique. *Trends in Cognitive Sciences*, 9, 21–25.
- Jaeger, T. (2008). Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language*, 59, 434–446.
- Jeannerod, M. (2001). Neural simulation of action: a unifying mechanism for motor cognition. *NeuroImage*, 14, 103–109.
- Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, 17, 187–245.
- Jeannerod, M., Arbib, M. A., Rizzolatti, G., & Sakata, H. (1995). Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends in Neurosciences*, 18, 314–320.
- Jeannerod, M. (2003). The mechanism of self-recognition in humans. *Behavioural Brain Research*, 142, 1–15.
- Jeannerod, M. (2006). *Motor cognition: What actions tell the self*. Oxford: Oxford University Press.

- Johnson-Frey, S. H., Maloof, F. R., Newman-Norlund, R., Farrer, C., Inati, S., & Grafton, S. T. (2003). Actions or hand-object interactions? Human inferior frontal cortex and action observation. *Neuron*, 39(6), 1053–1058.
- Johnson-Laird, P. N., & Stevenson, R. (1970). Memory for syntax. *Nature*, 227, 412.
- Johnson-Laird, P. N. (1980). Mental Models in Cognitive Science. *Cognitive Science*, 4, 71–115.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a Cognitive Science of Language, Inference and Consciousness*. Cambridge: Cambridge University Press.
- Joliceœur, P. (1999). Concurrent response-selection demands modulate the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1091–1113.
- Jonas, M., Biermann-Ruben, K., Kessler, K., Lange, R., Bäumer, T., Siebner, H. R., ... Münchau, A. (2007). Observation of a finger or an object movement primes imitative responses differentially. *Experimental Brain Research*, 177, 255–265.
- Kable, J. W., Kan, I. P., Wilson, A., Thompson-Schill, S. L., & Chatterjee, A. (2005). Conceptual representations of action in the lateral temporal cortex. *Journal of Cognitive Neuroscience*, 17, 1855–1870.
- Kable, J. W., Lease-Spellmeyer, J., & Chatterjee, A. (2002). Neural substrates of action event knowledge. *Journal of Cognitive Neuroscience*, 14, 795–805.
- Kaschak, M. P., & Borreggine, K. L. (2008). Temporal dynamics of the action-sentence compatibility effect. *Quarterly Journal of Experimental Psychology*, 61, 883–895.
- Kaschak, M. P., & Glenberg, A. M. (2000). Constructing meaning: The role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory and Language*, 43, 508–529.
- Kaschak, M. P., Madden, C. J., Theriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A., & Zwaan, R. A. (2005). Perception of motion affects language processing. *Cognition*, 94, 79–89.
- Kaschak, M. P., Zwaan, R. A., Aveyard, M., & Yaxley, R. H. (2006). Perception of auditory motion affects language processing. *Cognitive Science*, 30, 733–744.
- Kelley, W. M., Macrae, C. N., Wyland, C. L., Caglar, S., Inati, S., & Heatherton, T. F. (2002). Finding the self? An event-related fMRI study. *Journal of Cognitive Neuroscience*, 14, 785–794.
- Kemmerer, D., Rudrauf, D., Manzel, K., & Tranel, D. (2012). Behavioral patterns and lesion sites associated with impaired processing of lexical and conceptual knowledge of actions. *Cortex*, 48, 826–848.

- Kerzel, D., & Bekkering, H. (2000). Motor activation From visible speech : Evidence from stimulus response compatibility. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 634–647.
- Kessler, K., & Rutherford, H. (2010). The two forms of visuo-spatial perspective taking are differently embodied and subserve different spatial prepositions. *Frontiers in Psychology*, 1, 1–12.
- Keysers, C., Kohler, E., Umiltà, M. A., Nanetti, L., Fogassi, L., & Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research*, 153, 628–636.
- Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: theoretical developments, current evidence and future directions. *Cortex*, 48, 805–825.
- Kilner, J., Hamilton, A. F. D. C., & Blakemore, S.-J. (2007). Interference effect of observed human movement on action is due to velocity profile of biological motion. *Social Neuroscience*, 2, 158–166.
- Kilner, J. M., Paulignan, Y., & Blakemore, S.-J. (2003). An Interference Effect of Observed Biological Movement on Action. *Current Biology*, 13, 522–525.
- Kilner, J. M., Neal, A., Weiskopf, N., Friston, K. J., & Frith, C. D. (2009). Evidence of mirror neurons in human inferior frontal gyrus. *The Journal of Neuroscience*, 29, 10153–10159.
- Kilner, J. M., Vargas, C., Duval, S., Blakemore, S.-J., & Sirigu, A. (2004). Motor activation prior to observation of a predicted movement. *Nature Neuroscience*, 7, 1299–1301.
- Klatzky, R., Pellegrino, J., McCloskey, B., & Doherty, S. (1989). Can You Squeeze a Tomato? The Role of Motor Representations in Semantic Sensibility Judgments. *Journal of Memory and Language*, 28, 56–77.
- Klein, S. B., & Kihlstrom, J. F. (1986). Elaboration, organization, and the self-reference effect in memory. *Journal of Experimental Psychology: General*, 115, 26.
- Knoblich, G., & Flach, R. (2001). Predicting the effects of actions: Interactions of perception and action. *Psychological Science*, 12, 467–472.
- Knoblich, G., & Prinz, W. (2001). Recognition of self-generated actions from kinematic displays of drawing. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 456–465.
- Knoblich, G., & Repp, B. H. (2009). Inferring agency from sound. *Cognition*, 111, 248–262.
- Knoblich, G., & Flach, R. (2003). Action identity: Evidence from self-recognition, prediction, and coordination. *Consciousness and Cognition*, 12, 620–632.
- Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: Action representation in mirror neurons. *Science*, 297, 846–848.

- Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 47–60.
- Kunde, W., & Wühr, P. (2004). Actions blind to conceptually overlapping stimuli. *Psychological Research*, 68, 199–207.
- Kutas, M., & Federmeier, K. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, 4, 463–470.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato’s problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 1, 211–240.
- Lesage, E., Morgan, B. E., Olson, A. C., Meyer, A. S., & Miall, R. C. (2012). Cerebellar rTMS disrupts predictive language processing. *Current Biology*, 22, R794–R795.
- Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–38; discussion 38–75.
- Levinson, S. C. (1996). Frames of reference and Molyneux’s question: Cross-linguistic evidence. In P. Bloom, M. Peterson, L. Nadel, & M. Garrett (Eds.), *Language and space* (pp. 109–169). Cambridge, MA: MIT Press.
- Levinson, S. C. (2003). *Space and Language in Cognition: Explorations in Cognitive Diversity*. Cambridge: Cambridge University Press.
- Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21, 1–36.
- Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74, 431–461.
- Liepelt, R., & Brass, M. (2010). Top-down modulation of motor priming by belief about animacy. *Experimental Psychology*, 57, 221–227.
- Lindemann, O., Stenneken, P., van Schie, H. T., & Bekkering, H. (2006). Semantic activation in action planning. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 633–643.
- Lindsay, S., Scheepers, C., & Kamide, Y. (2013). To dash or to dawdle: Verb-associated speed of motion influences eye movements during spoken sentence comprehension. *PLOS ONE*, 8, e67187.
- Loula, F., Prasad, S., Harber, K., & Shiffrar, M. (2005). Recognizing people from their movement. *Journal of Experimental Psychology: Human Perception and Performance*, 31, 210–220.
- Louwerse, M. M., & Jeuniaux, P. (2008). Language comprehension is both embodied and symbolic. In M. de Vega, A. Glenberg, & A. C. Graesser (Eds.), *Symbols, embodiment, and meaning* (pp. 309–326). Oxford: Oxford University Press.

- Lyons, I. M., Mattarella-Micke, A., Cieslak, M., Nusbaum, H. C., Small, S. L., & Beilock, S. L. (2010). The role of personal experience in the neural processing of action-related language. *Brain and Language*, *112*, 214–222.
- Macdonald, M. C., Paerlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 673–703.
- Maeda, F., Kleiner-Fisman, G., & Pascual-Leone, A. (2002). Motor facilitation while observing hand actions : Specificity of the effect and role of observer's orientation. *Journal of Neurophysiology*, *87*, 1329–1335.
- Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology*, *102*, 59–70.
- Mainwaring, S. D., Tversky, B., Ohgishi, M., & Schiano, D. J. (2009). Descriptions of Simple Spatial Scenes in English and Japanese. *Spatial Cognition & Computation*, *3*, 3–42.
- Marino, B. F. M., Gallese, V., Buccino, G., & Riggio, L. (2012). Language sensorimotor specificity modulates the motor system. *Cortex*, *48*, 849–856.
- Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: Structure and processes. *Current Opinion in Neurobiology*, *11*, 194–201.
- Martin, A., Wiggs, C., Ungerleider, L., & Haxby, J. (1996). Neural correlates of category specific knowledge. *Nature*, *379*, 649–652.
- Marzoli, D., Mitaritonna, A., Moretto, F., Carluccio, P., & Tommasi, L. (2011). The handedness of imagined bodies in action and the role of perspective taking. *Brain and Cognition*, *75*, 51–59.
- Masson, M. E. J., Bub, D. N., & Newton-Taylor, M. (2008). Language-based access to gestural components of conceptual knowledge. *Quarterly Journal of Experimental Psychology*, *61*, 869–882.
- Matlock, T. (2004). Fictive motion as cognitive simulation. *Memory and Cognition*, *32*, 1389–1400.
- Mazzarella, E., Hamilton, A., Trojano, L., Mastromauro, B., & Conson, M. (2012). Observation of another's action but not eye gaze triggers allocentric visual perspective. *Quarterly Journal of Experimental Psychology*, *65*, 2447–2460.
- McCabe, D. P., & Castel, A. D. (2008). Seeing is believing: the effect of brain images on judgments of scientific reasoning. *Cognition*, *107*, 343–352.
- Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of age: a review of embodiment and the neuroscience of semantics. *Cortex*, *48*, 788–804.

- Miall, R. C., Stanley, J., Todhunter, S., Levick, C., Lindo, S., & Miall, J. D. (2006). Performing hand actions assists the visual discrimination of similar hand postures. *Neuropsychologia*, 44, 966–976.
- Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Networks*, 9, 1265–1279.
- Michelon, P., & Zacks, J. M. (2006). Two kinds of visual perspective taking. *Perception and Psychophysics*, 68, 327–337.
- Mirabella, G., Iaconelli, S., Spadacenta, S., Federico, P., & Gallese, V. (2012). Processing of hand-related verbs specifically affects the planning and execution of arm reaching movements. *PLOS ONE*, 7, e35403.
- Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current Biology*, 20, 750–756.
- Müsseler, J., & Hommel, B. (1997). Blindness to response-compatible stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 861–872.
- Müsseler, J., & Hommel, B. (1997). Detecting and identifying response-compatible stimuli. *Psychonomic Bulletin & Review*, 4, 125–129.
- Müsseler, Jochen, Steininger, S., & Wühr, P. (2001). Can actions affect perceptual processing? *Quarterly Journal of Experimental Psychology*, 37–41.
- Müsseler, Jochen, Wühr, P., & Prinz, W. (2000). Varying the response code in the blindness to response-compatible stimuli. *Visual Cognition*, 7, 743–767.
- Nazir, T. A, Boulenger, V., Roy, A., Silber, B., Jeannerod, M., & Paulignan, Y. (2008). Language-induced motor perturbations during the execution of a reaching movement. *Quarterly Journal of Experimental Psychology*, 61, 933–943.
- Neininger, B., & Pulvermüller, F. (2003). Word-category specific deficits after lesions in the right hemisphere. *Neuropsychologia*, 41, 53–70.
- Noppeney, U., Josephs, O., Kiebel, S., Friston, K. J., & Price, C. J. (2005). Action selectivity in parietal and temporal cortex. *Cognitive Brain Research*, 25, 641–649.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97–113.
- Oliveri, M., Finocchiaro, C., Shapiro, K., Gangitano, M., Caramazza, A., & Pascual-Leone, A. (2004). All talk and no action: a transcranial magnetic stimulation study of motor cortex activation during action word production. *Journal of Cognitive Neuroscience*, 16, 374–381.
- Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 17, 776.

- Page, M. P. A. (2005). What can't functional neuroimaging tell the cognitive psychologist? *Cortex*, 42, 428–443.
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, 45, 255.
- Papagno, C., Curti, R., Rizzo, S., Crippa, F., & Colombo, M. R. (2006). Is the right hemisphere involved in idiom comprehension? A neuropsychological study. *Neuropsychology*, 20, 598–606.
- Papeo, L., Corradi-Dell'Acqua, C., & Rumiati, R. I. (2011). "She" is not like "I": the tie between language and action is in our imagination. *Journal of Cognitive Neuroscience*, 23, 3939–3948.
- Papeo, L., Rumiati, R. I., Cecchetto, C., & Tomasino, B. (2012). On-line changing of thinking about words: The effect of cognitive context on neural responses to verb reading. *Journal of Cognitive Neuroscience*, 24, 2348–2362.
- Papeo, L., Vallesi, A., Isaja, A., & Rumiati, R. I. (2009). Effects of TMS on different stages of motor and non-motor verb processing in the primary motor cortex. *PLOS ONE*, 4, e4508.
- Parsons, L. M. (1994). Temporal and kinematic properties of motor behavior reflected in mentally simulated action. *Journal of Experimental Psychology. Human Perception and Performance*, 20, 709–730.
- Parsons, L. M. (1987). Imagined spatial transformations of one's hand and feet. *Cognitive Psychology*, 24, 178–241.
- Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, 116, 220–244.
- Pecher, D., & Zwaan, R. A. (Eds.). (2005). *Grounding cognition: The role of perception and action in memory, language, and thinking*. Cambridge: Cambridge University Press.
- Pecher, D., van Dantzig, S., Zwaan, R. A., & Zeelenberg, R. (2009). Language comprehenders retain implied shape and orientation of objects. *Quarterly Journal of Experimental Psychology*, 62, 1108–1114.
- Pennfield, W., & Rasmussen, T. (1950). *The cerebral cortex of man: A clinical study of localisation of function*. New York: Macmillan.
- Péran, P., Cardebat, D., Cherubini, A., Piras, F., Luccichenti, G., Peppe, A., ... Sabatini, U. (2009). Object naming and action-verb generation in Parkinson's disease: a fMRI study. *Cortex*, 45, 960–971.
- Péran, P., Démonet, J.-F., Cherubini, A., Cardebat, D., Caltagirone, C., & Sabatini, U. (2010). Mental representations of action: the neural correlates of the verbal and motor components. *Brain research*, 1328, 89–103.

- Péran, P., Rascol, O., Celsis, P., Nespoulous, J., & Dubois, B. (2003). Deficit of verb generation in nondemented patients with Parkinson's disease. *Movement Disorders*, *18*, 150–156.
- Perani, D., Fazio, F., Borghese, N. A., Tettamanti, M., Ferrari, S., Decety, J., & Gilardi, M. C. (2001). Different brain correlates for watching real and virtual hand actions. *NeuroImage*, *14*, 749–758.
- Petroni, A., Baguear, F., & Della-Maggiore, V. (2010). Motor resonance may originate from sensorimotor experience. *Journal of Neurophysiology*, *104*, 1867–1871.
- Pezzulo, G., & Castelfranchi, C. (2009). Thinking as the control of imagination: a conceptual framework for goal-directed systems. *Psychological Research*, *73*, 559–577.
- Pickering, M. J., & Ferreira, V. S. (2009). Structural priming: A critical review. *Psychological Bulletin*, *134*, 427–459.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*, 169–90.
- Pickering, M. J., & Garrod, S. (2006). Alignment as the basis for successful communication. *Research on Language and Computation*, *4*, 203–228.
- Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, *11*, 105–110.
- Pickering, M. J., & Garrod, S. (2009). Prediction and embodiment in dialogue. *European Journal of Social Psychology*, *39*, 1162–1168.
- Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*, 329–347.
- Pickering, M. J., McLean, J. F., & Gambi, C. (2012). Do addressees adopt the perspective of the speaker? *Acta Psychologica*, *141*, 261–269.
- Posner, M. I., Boies, S. J., Eichelman, W. H., & Taylor, R. L. (1969). Retention of visual and name codes of single letters. *Journal of Experimental Psychology*, *79*, 1–16.
- Postle, N., McMahon, K. L., Ashton, R., Meredith, M., & de Zubicaray, G. I. (2008). Action word meaning representations in cytoarchitectonically defined primary and premotor cortices. *NeuroImage*, *43*, 634–644.
- Prather, J. F., Peters, S., Nowicki, S., & Mooney, R. (2008). Precise auditory-vocal mirroring in neurons for learned vocal communication. *Nature*, *451*, 305–310.
- Press, C., Bird, G., Flach, R., & Heyes, C. (2005). Robotic movement elicits automatic imitation. *Cognitive Brain Research*, *25*, 632–640.

- Press, C., Gillmeister, H., & Heyes, C. (2006). Bottom-up, not top-down, modulation of imitation by human and robotic models. *The European Journal of Neuroscience*, 24, 2415–2419.
- Press, C., Gillmeister, H., & Heyes, C. (2007). Sensorimotor experience enhances automatic imitation of robotic action. *Proceedings. Biological sciences / The Royal Society*, 274(1625), 2509–2514.
- Prinz, W. (2006). What re-enactment earns us. *Cortex*, 42, 515–517.
- Prinz, W. (1987). Ideo-motor action. In H. Heuer & E. Sanders (Eds.), *Perspectives on perception and action* (pp. 47–76). Hillsdale, NJ: Erlbaum.
- Prinz, W. (1997). Perception and Action Planning. *European Journal of Cognitive Psychology*, 9, 129–154.
- Prinz, W., & Hommel, B. (Eds.). (2002). *Common mechanisms in perception and action: Attention and performance XIV*. Oxford: Oxford University Press.
- Pulvermüller, F., Härle, M., & Hummel, F. (2000). Neurophysiological distinction of verb categories. *Neuroreport*, 11, 2789–2793.
- Pulvermüller, F., Härle, M., & Hummel, F. (2001). Walking or talking? Behavioral and neurophysiological correlates of action verb processing. *Brain and Language*, 78, 143–168.
- Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences*, 22, 253–336.
- Pulvermüller, F. (2001). Brain reflections of words and their meaning. *Trends in Cognitive Sciences*, 5, 517–524.
- Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, 6, 576–582.
- Pulvermüller, F., & Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nature reviews Neuroscience*, 11, 351–360.
- Pulvermüller, F., Hauk, O., Nikulin, V., & Ilmoniemi, R. J. (2005). Functional links between motor and language systems. *The European Journal of Neuroscience*, 21, 793–797.
- Pulvermüller, F., Huss, M., Kherif, F., Moscoso del Prado Martin, F., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 7865–7870.
- Pulvermüller, F., Shtyrov, Y., & Hauk, O. (2009). Understanding in an instant: Neurophysiological evidence for mechanistic language circuits in the brain. *Brain and Language*, 110, 81–94.
- Pulvermüller, F., Shtyrov, Y., & Ilmoniemi, R. (2005). Brain signatures of meaning access in action word recognition. *Journal of Cognitive Neuroscience*, 17, 884–892.

- Pylyshyn, Z. W. (1984). *Computation and cognition*. Cambridge, MA: MIT Press.
- R Development Core Team. (2011). R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from <http://www.r-project.org/>
- Raffray, C. N., Pickering, M. J., & Branigan, H. P. (2007). Priming the interpretation of noun–noun combinations. *Journal of Memory and Language*, 57, 380–395.
- Ransil, B. J., & Schachter, S. C. (1994). Test-retest reliability of the Edinburgh Handedness Inventory and Global Handedness preference measurements, and their correlation. *Perceptual and Motor Skills*, 79, 1355–1372.
- Raposo, A., Moss, H. E., Stamatakis, E. A., & Tyler, L. K. (2009). Modulation of motor and premotor cortices by actions, action words and action sentences. *Neuropsychologia*, 47, 388–396.
- Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of Perceptual Fluency on Affective Judgments. *Psychological Science*, 9, 45–48.
- Repp, B. H., & Keller, P. E. (2010). Self versus other in piano performance: Detectability of timing perturbations depends on personal playing style. *Experimental Brain Research*, 202, 101–110.
- Repp, B. H., & Knoblich, G. (2004). Perceiving action identity: how pianists recognize their own performances. *Psychological Science*, 15, 604–609.
- Repp, B. H., & Knoblich, G. (2007). Action can affect auditory perception. *Psychological Science*, 18, 6–7.
- Reuschmeyer, S.-A., & Bekkering, H. (2013). Embodied lexical representations: Flexible tools for predicting the future. In Y. Coello & A. Bartolo (Eds.), *Language and action in cognitive neuroscience* (pp. 111–124). Hove: Psychology Press.
- Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Brain Research*, 3, 131–141.
- Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, 21, 188–914.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169–192.
- Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (1999). Resonance behaviours and mirror neurons. *Archives Italiennes de Biologie*, 137, 85–100.
- Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, 2, 1–10.
- Rizzolatti, G., & Luppino, G. (2001). The cortical motor system. *Neuron*, 31, 889–901.

- Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature reviews Neuroscience*, *11*, 264–274.
- Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology*, *35*, 677.
- Rubichi, S., & Nicoletti, R. (2006). The Simon effect and handedness: evidence for a dominant-hand attentional bias in spatial coding. *Perception & Psychophysics*, *68*, 1059–1069.
- Ruby, P., & Decety, J. (2001). Effect of subjective perspective taking during simulation of action: A PET investigation of agency. *Nature Neuroscience*, *4*, 546–550.
- Ruby, P., & Decety, J. (2004). How would you feel versus how do you think she would feel? A neuroimaging study of perspective-taking with social emotions. *Journal of Cognitive Neuroscience*, *16*, 988–999.
- Rueschemeyer, S.-A., Lindemann, O., van Rooij, D., van Dam, W., & Bekkering, H. (2010). Effects of intentional motor actions on embodied language processing. *Experimental Psychology*, *57*, 260–266.
- Rueschemeyer, S.-A., Pfeiffer, C., & Bekkering, H. (2010). Body schematics: on the role of the body schema in embodied lexical-semantic representations. *Neuropsychologia*, *48*, 774–781.
- Rueschemeyer, S.-A., van Rooij, D., Lindemann, O., Willems, R. M., & Bekkering, H. (2010). The function of words: Distinct neural correlates for words denoting differently manipulable objects. *Journal of Cognitive Neuroscience*, *22*, 1844–1851.
- Saccuman, M. C., Cappa, S. F., Bates, E. A., Arevalo, A., Della Rosa, P., Danna, M., & Perani, D. (2006). The impact of semantic reference on word class: An fMRI study of action and object naming. *NeuroImage*, *32*, 1865–1878.
- Sachs, J. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics*, *2*, 437–442.
- Sartori, L., Begliomini, C., & Castiello, U. (2013). Motor resonance in left- and right-handers: evidence for effector-independent motor representations. *Frontiers in Human Neuroscience*, *7*, 33.
- Sato, M., & Bergen, B. K. (2013). The case of the missing pronouns: Does mentally simulated perspective play a functional role in the comprehension of person. *Cognition*, *127*, 361–374.
- Sato, M., Mengarelli, M., Riggio, L., Gallese, V., & Buccino, G. (2008). Task related modulation of the motor system during language processing. *Brain and Language*, *105*, 83–90.

- Schieber, M. H. (2001). Constraints on somatotopic organization in the primary motor cortex constraints on somatotopic organization in the primary motor cortex. *Journal of Neurophysiology*, 86, 2125–2143.
- Schober, M. (1993). Spatial perspective-taking in conversation. *Cognition*, 47, 1–24.
- Schober, M. (1995). Speakers, addressees, and frames of reference: Whose effort is minimized in coversations about locations? *Discourse Processes*, 20, 219–247.
- Schober, M. (2009). Spatial dialogue between partners with mismatched abilities. In K. Coventry, T. Tenbrink, & J. Bateman (Eds.), *Spatial language and dialogue* (pp. 23–39). Oxford: Oxford University Press.
- Schober, M., & Brennan, S. E. (2003). Processes of interactive spoken discourse: The role of the partner. In A. C. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds.), *Handbook of discourse processes* (pp. 123–164). Mahwah, NJ: Lawrence Erlbaum.
- Schubö, A., Aschersleben, G., & Prinz, W. (2001). Interactions between perception and action in a reaction task with overlapping S-R assignments. *Psychological Research*, 65, 145–157.
- Schubö, A., Prinz, W., & Aschersleben, G. (2004). Perceiving while acting: Action affects perception. *Psychological Research*, 68, 208–15.
- Schütz-Bosbach, S., Mancini, B., Aglioti, S. M., & Haggard, P. (2006). Self and other in the human motor system. *Current Biology*, 16, 1830–1834.
- Schütz-Bosbach, S., & Prinz, W. (2007). Perceptual resonance: action-induced modulation of perception. *Trends in Cognitive Sciences*, 11, 349–355.
- Scorolli, C., Borghi, A. M., & Glenberg, A. (2009). Language-induced motor activity in bi-manual object lifting. *Experimental Brain Research*, 193, 43–53.
- Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences*, 10, 70–76.
- Secora, K., & Emmorey, K. (2013). The Action-Sentence Compatibility Effect in American Sign Language. In *Poster presented at the 26th Annual CUNY Conference on Human Sentence Processing*. University of South Carolina, Columbia.
- Sevdalis, V., & Keller, P. E. (2010). Cues for self-recognition in point-light displays of actions performed in synchrony with music. *Consciousness and Cognition*, 19, 617–626.
- Seymour, P. H. K. (1969). Pictorial coding of verbal descriptions. *Quarterly Journal of Experimental Psychology*, (December 1969), 37–41.
- Shapiro, L. (2010). *Embodied cognition*. Oxford: Taylor & Francis.
- Shintel, H., Nusbaum, H. C., & Okrent, A. (2006). Analog acoustic expression in speech communication. *Journal of Memory and Language*, 55, 167–177.

- Shockley, K., Santana, M.-V., & Fowler, C. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation.pdf. *Journal of Experimental Psychology: Human Perception and Performance*, 29, 326–332.
- Shtyrov, Y., Hauk, O., & Pulvermuller, F. (2004). Distributed neuronal networks for encoding category-specific semantic information : the mismatch negativity to action words. *European Journal of Neuroscience*, 19, 1083–1092.
- Simmons, W. K., & Barsalou, L. W. (2003). The similarity-in-topography principle: reconciling theories of conceptual deficits. *Cognitive Neuropsychology*, 20, 451–486.
- Sirigu, A., Daprati, E., Pradat-Diehl, P., Franck, N., & Jeannerod, M. (1999). Perception of self-generated movement following left parietal lesion. *Brain*, 122, 1867–1874.
- Small, J. A., Kemper, S., & Lyons, K. (1997). Sentence comprehension in Alzheimer's disease: Effects of grammatical complexity, speech rate, and repetition. *Psychology and Aging*, 12, 3–11.
- Southgate, V., Johnson, M. H., Osborne, T., & Csibra, G. (2009). Predictive motor activation during action observation in human infants. *Biology Letters*, 5, 769–772.
- Stanfield, R., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, 12, 153–156.
- Stanley, J., Gowen, E., & Miall, R. C. (2007). Effects of agency on movement interference during observation of a moving dot stimulus. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 915–926.
- Stevanovski, B., Oriet, C., & Jolicoeur, P. (2002). Blinded by headlights. *Canadian Journal of Experimental Psychology*, 56, 65–74.
- Stoet, G., & Hommel, B. (2002). Interaction between feature binding in perception and action. (Wolfgang Prinz & B. Hommel, Eds.) *Common mechanisms in perception and action*. Oxford: Oxford University Press.
- Strafella, A. P., & Paus, T. (2000). Modulation of cortical excitability during action observation: A transcranial magnetic stimulation study. *Neuroreport*, 11, 2289–2292.
- Stürmer, B., Aschersleben, G., & Prinz, W. (2000). Correspondence effects with manual gestures and postures: A study of imitation. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 1746–1759.
- Sturt, P., Sanford, A. J., Stewart, A., & Dawydiak, E. (2004). Linguistic focus and good-enough representations: an application of the change-detection paradigm. *Psychonomic Bulletin & Review*, 11, 882–888.
- Summers, J. J., Semjen, A., Carson, R. G., & Thomas, J. (1995). Going around in circles: The dynamics of bimanual circling. In D. J. Glencross & J. P. Piek (Eds.), *Motor control and sensory motor integration* (pp. 231–253). Amsterdam: Elsevier.

- Symes, E., Ellis, R., & Tucker, M. (2005). Dissociating object-based and space-based affordances. *Visual Cognition*, 12, 1337–1361.
- Symes, E., Tucker, M., Ellis, R., Vainio, L., & Ottoboni, G. (2008). Grasp preparation improves change detection for congruent objects. *Journal of Experimental Psychology. Human Perception and Performance*, 34, 854–871.
- Tai, Y. F., Scherfler, C., Brooks, D. J., Sawamoto, N., & Castiello, U. (2004). The human premotor cortex Is “mirror” only for biological actions. *Current Biology*, 14, 117–120.
- Taylor, L. J., Lev-Ari, S., & Zwaan, R. A. (2008). Inferences about action engage action systems. *Brain and Language*, 107, 62–67.
- Taylor, L. J., & Zwaan, R. A. (2008). Motor resonance and linguistic focus. *Quarterly Journal of Experimental Psychology*, 61, 896–904.
- Taylor, L. J., & Zwaan, R. A. (2009). Action in cognition: The case of language. *Language and Cognition*, 1, 45–58.
- Taylor, L. J., & Zwaan, R. A. (2013). Fault-tolerant comprehension. In Y. Coello & A. Bartolo (Eds.), *Language and action in cognitive neuroscience* (pp. 145–158). Hove: Psychology Press.
- Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., ... Perani, D. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, 17, 273–281.
- Thach, W. T. (1998). What is the role of the cerebellum in motor learning and cognition?. *Trends in Cognitive Sciences*, 2, 331–337.
- Tomasino, B., & Rumiati, R. I. (2013). At the mercy of strategies: the role of motor representations in language understanding. *Frontiers in Psychology*, 4, 27.
- Tomasino, B., Werner, C. J., Weiss, P. H., & Fink, G. R. (2007). Stimulus properties matter more than perspective: an fMRI study of mental imagery and silent reading of action phrases. *NeuroImage*, 36, 128–141.
- Triggs, W. J., Subramaniam, B., & Rossi, F. (1999). Hand preference and transcranial magnetic stimulation asymmetry of cortical motor representation. *Brain Research*, 835, 324–329.
- Tucker, M., & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology. Human Perception and Performance*, 24, 830–846.
- Tucker, M., & Ellis, R. (2001). The potentiation of grasp types during visual object categorization. *Visual Cognition*, 8, 769–800.
- Tucker, M., & Ellis, R. (2004). Action priming by briefly presented objects. *Acta Psychologica*, 116, 185–203.

- Tversky, B. (1991). Spatial mental models. *The Psychology of Learning and Motivation: Advances in Research and Theory*, 27, 109–145.
- Tversky, B., & Hard, B. M. (2009). Embodied and disembodied cognition: spatial perspective-taking. *Cognition*, 110, 124–129.
- Tye-Murray, N., Spehar, B. P., Myerson, J., Hale, S., & Sommers, M. S. (2013). Reading your own lips: common-coding theory and visual speech perception. *Psychonomic Bulletin & Review*, 20, 115–119.
- Umiltà, M. A., Escola, L., Intskirveli, I., Grammont, F., Rochat, M., Caruana, F., ... Rizzolatti, G. (2008). When pliers become fingers in the monkey motor system. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 2209–2213.
- Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001). I know what you are doing : A neurophysiological study. *Neuron*, 31, 155–165.
- Van Dam, W. O., Rueschemeyer, S.-A., Lindemann, O., & Bekkering, H. (2010). Context effects in embodied lexical-semantic processing. *Frontiers in Psychology*, 1, 150.
- Van Dam, W. O., van Dijk, M., Bekkering, H., & Rueschemeyer, S.-A. (2012). Flexibility in embodied lexical-semantic representations. *Human Brain Mapping*, 33, 2322–2333.
- Van der Lely, H. K. J. (1996). Specifically language impaired and normally developing children: Verbal passive vs. adjectival passive sentence interpretation. *Lingua*, 98, 243–272.
- Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Van Elk, M., van Schie, H. T., & Bekkering, H. (2009). Short-term action intentions overrule long-term semantic knowledge. *Cognition*, 111, 72–83.
- Van Elk, M., & Blanke, O. (2011). The relation between body semantics and spatial body representations. *Acta Psychologica*, 138, 347–358.
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.
- Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). Representing the meanings of object and action words: the featural and unitary semantic space hypothesis. *Cognitive Psychology*, 48, 422–488.
- Vigliocco, G., Warren, J., Siri, S., Arciuli, J., Scott, S., & Wise, R. (2006). The role of semantics and grammatical class in the neural representation of words. *Cerebral Cortex*, 16, 1790–1796.
- Vogt, S., Taylor, P., & Hopkins, B. (2003). Visuomotor priming by pictures of hand postures: Perspective matters. *Neuropsychologia*, 41, 941–951.

- Wales, R. J., & Grieve, R. (1969). What Is so difficult about negation? *Perception & Psychophysics*, 6, 327–332.
- Watson, M. E., Pickering, M. J., & Branigan, H. P. (1994). Alignment of Reference Frames in Dialogue. In *Proceedings of the 26th Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum.
- Weisberg, D. S., Keil, F. C., Goodstein, J., Rawson, E., & Gray, J. R. (2008). The seductive allure of neuroscience explanations. *Journal of Cognitive Neuroscience*, 20, 470–477.
- Willems, R. M., & Francken, J. C. (2012). Embodied cognition: taking the next step. *Frontiers in Psychology*, 3, 582.
- Willems, R. M., & Hagoort, P. (2007). Neural evidence for the interplay between language, gesture, and action: a review. *Brain and Language*, 101, 278–289.
- Willems, R. M., Hagoort, P., & Casasanto, D. (2010). Body-specific representations of action verbs: neural evidence from right- and left-handers. *Psychological Science*, 21, 67–74.
- Willems, R. M., Labruna, L., D’Esposito, M., Ivry, R., & Casasanto, D. (2011). A functional role for the motor system in language understanding: evidence from theta-burst transcranial magnetic stimulation. *Psychological Science*, 22, 849–854.
- Willems, R. M., Toni, I., Hagoort, P., & Casasanto, D. (2009). Body-specific motor imagery of hand actions: neural evidence from right- and left-handers. *Frontiers in Human Neuroscience*, 3, 1–9.
- Willems, R. M., Toni, I., Hagoort, P., & Casasanto, D. (2010). Neural dissociations between action verb understanding and motor imagery. *Journal of Cognitive Neuroscience*, 22, 2387–2400.
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9, 625–636.
- Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, 131, 460–473.
- Witt, J. K., Kemmerer, D., Linkenauger, S. a, & Culham, J. (2010). A functional role for motor simulation in identifying tools. *Psychological Science*, 21, 1215–1219.
- Wohlschläger, A. (2000). Visual motion priming by invisible actions. *Vision Research*, 40, 925–930.
- Wolpert, D M. (1997). Computational approaches to motor control. *Trends in Cognitive Sciences*, 1, 209–216.
- Wolpert, D.M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology*, 11, R729–732.
- Wolpert, D.M., Ghahramani, Z., & Jordan, M. (1995). An internal model for sensorimotor integration. *Science*, 269, 1880–1882.

- Wolpert, D. M., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, *11*, 1317-1329.
- Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences*, *2*, 338-347.
- Wraga, M., Thompson, W. L., Alpert, N. M., & Kosslyn, S. M. (2003). Implicit transfer of motor strategies in mental rotation. *Brain and Cognition*, *52*, 135-143.
- Wu, L., & Barsalou, L. W. (2009). Perceptual simulation in conceptual combination: evidence from property generation. *Acta Psychologica*, *132*, 173-189.
- Wykowska, A., Schubö, A., & Hommel, B. (2009). How you move is what you see: Action planning biases selection in visual search. *Journal of Experimental Psychology. Human Perception and Performance*, *35*, 1755-1769.
- Yoon, S. O., Koh, S., & Brown-Schmidt, S. (2012). Influence of perspective and goals on reference production in conversation. *Psychonomic Bulletin & Review*, *19*, 699-707.
- Zwaan, R. A., Taylor, L. J., & de Boer, M. (2010). Motor resonance as a function of narrative time: further tests of the linguistic focus hypothesis. *Brain and Language*, *112*, 143-149.
- Zwaan, R. A. (2008). Time in language, situation models, and mental simulations. *Language Learning*, *58*, 13-26.
- Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The Construction of Situation Models in Narrative Comprehension: An Event-Indexing Model. *Psychological Science*, *6*, 292-297.
- Zwaan, R. A., & Madden, C. J. (2005). Embodied sentence comprehension. In D. Pecher & R. A. Zwaan (Eds.), *Grounding cognition: The role of action in memory, language and thinking*. (pp. 224-245). Cambridge: Cambridge University Press.
- Zwaan, R. A., & Radvansky, G. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162-85.
- Zwaan, R. A., & Ross, B. H. (2004). The immersed experiencer: Toward an embodied theory of language. In *The Psychology of Learning and Motivation: Advances in Research and Theory* (pp. 35-62). San Diego, CA: Academic Press.
- Zwaan, R. A., Stanfield, R., & Yaxley, R. H. (2002). Language Comprehenders Mentally Represent the Shapes of Objects. *Psychological Science*, *13*, 168-171.
- Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: motor resonance in language comprehension. *Journal of Experimental Psychology: General*, *135*, 1-11.
- Zwicker, J. (2009). Agency attribution and visuospatial perspective taking. *Psychonomic Bulletin & Review*, *16*, 1089-1093.

- Zwicker, J., Grosjean, M., & Prinz, W. (2008). A contrast effect between the concurrent production and perception of movement directions. *Visual Cognition*, 16, 953–978.
- Zwicker, J., Grosjean, M., & Prinz, W. (2010). On interference effects in concurrent perception and action. *Psychological Research*, 74, 152–171.

APPENDIX A: THE EDINBURGH HANDEDNESS

INVENTORY (from Oldfield, 1971)

Participant number: _____

Please indicate with a check (✓) your preference in using your left or right hand in the following tasks.

Where the preference is so strong you would never use the other hand, unless absolutely forced to, put two checks (✓✓).

If you are indifferent, put one check in each column (✓ | ✓).

Some of the activities require both hands. In these cases, the part of the task or object for which hand preference is wanted is indicated in parentheses.

| Task / Object | Left Hand | Right Hand |
|---|----------------------|------------|
| 1. Writing | | |
| 2. Drawing | | |
| 3. Throwing | | |
| 4. Scissors | | |
| 5. Toothbrush | | |
| 6. Knife (without fork) | | |
| 7. Spoon | | |
| 8. Broom (upper hand) | | |
| 9. Striking a Match (match) | | |
| 10. Opening a Box (lid) | | |
| Total checks: | LH = | RH = |
| Cumulative Total | CT = LH + RH = | |
| Difference | D = RH – LH = | |
| Result | R = (D / CT) × 100 = | |
| Interpretation: (Left Handed: R < -40) (Ambidextrous: -40 ≤ R ≤ +40) (Right Handed: R > +40) | | |

APPENDIX B: MATERIALS IN EXPERIMENTS 1-6

SENTENCES

Experiment 1 used only first-person sentences. Experiments 2a – 6 used both first-person and third-person sentences. In Experiment 1, a sentence was matched with an image showing the matching action on critical trials, and with an image showing the object mismatch or action mismatch action on filler trials. In Experiments 2a – 6, a sentence was matched with a pair of images showing the matching action on critical trials, and with one image showing the matching action and one image showing an object mismatch or action mismatch on filler trials.

| Sentence | Matching action | Object mismatch action | Action mismatch action |
|--|--------------------|------------------------|------------------------|
| <i>I am/ He is slicing the bread</i> | Slice bread | | Butter bread |
| <i>I am/ He is slicing the tomato</i> | Slice tomato | Slice bread | |
| <i>I am/ He is buttering the bread</i> | Butter bread | | Cut bread |
| <i>I am/ He is examining the paper</i> | Examine paper | | Staple paper |
| <i>I am/ He is glueing the envelope</i> | Glue envelope | | Write on envelope |
| <i>I am/ He is glueing the paper</i> | Glue paper | | Examine paper |
| <i>I am/ He is cutting the paper</i> | Cut paper | Cut plant | |
| <i>I am/ He is cutting the plant</i> | Cut plant | | Rake plant |
| <i>I am/ He is holepunching the notebook</i> | Holepunch notebook | | Write in notebook |
| <i>I am/ He is holepunching the paper</i> | Holepunch paper | Holepunch notebook | |
| <i>I am/ He is ironing the cloth</i> | Iron cloth | Iron shirt | |
| <i>I am/ He is ironing the shirt</i> | Iron shirt | | Sew shirt |
| <i>I am/ He is lighting the candle</i> | Light candle | Light cigarette | |
| <i>I am/ He is lighting the cigarette</i> | Light cigarette | Light candle | |
| <i>I am/ He is opening the bottle</i> | Open bottle | Open can | |
| <i>I am/ He is opening the can</i> | Open can | Open bottle | |
| <i>I am/ He is peeling the carrot</i> | Peel carrot | Peel potato | |
| <i>I am/ He is peeling the potato</i> | Peel potato | Peel carrot | |
| <i>I am/ He is picking up the gherkin</i> | Pick up gherkin | Pick up olive | |
| <i>I am/ He is picking up the olive</i> | Pick up olive | | Scoop olive |
| <i>I am/ He is scooping up the olive</i> | Scoop up olive | Scoop yoghurt | |
| <i>I am/ He is scooping up the yoghurt</i> | Scoop yoghurt | Scoop olive | |
| <i>I am/ He is sewing the cloth</i> | Sew cloth | | Iron cloth |
| <i>I am/ He is sewing the shirt</i> | Sew shirt | | Scrub shirt |
| <i>I am/ He is scrubbing the pan</i> | Scrub pan | | Wipe pan |
| <i>I am/ He is scrubbing the shirt</i> | Scrub shirt | | Iron shirt |
| <i>I am/ He is stapling the paper</i> | Staple paper | | Cut paper |
| <i>I am/ He is stapling the envelope</i> | Staple envelope | | Glue envelope |
| <i>I am/ He is wiping the glass</i> | Wipe glass | Wipe pan | |
| <i>I am/ He is wiping the pan</i> | Wipe pan | Wipe glass | |
| <i>I am/ He is writing on the envelope</i> | Write on envelope | Write in notebook | |
| <i>I am/ He is writing in the notebook</i> | Write in notebook | Write in envelope | |

IMAGES

Experiment 1 used only internally orientated images. Experiments 2a – 6 used both internally and externally orientated images.

EXTERNAL IMAGES

INTERNAL IMAGES



EXTERNAL IMAGES



INTERNAL IMAGES



EXTERNAL IMAGES



INTERNAL IMAGES



EXTERNAL IMAGES



INTERNAL IMAGES



EXTERNAL IMAGES



INTERNAL IMAGES



EXTERNAL IMAGES

INTERNAL IMAGES



APPENDIX C: MATERIALS IN EXPERIMENTS 7 - 9

POKE sentences

I am buzzing the guard on the intercom
I am calling the lift down to the lobby
I am cancelling the print job to the printer
I am choosing a mars bar from the vending machine
I am clicking on the link to the website
I am dabbing some lipbalm on my lips
I am dialling 999 for an ambulance
I am dividing up the bill on the calculator
I am drawing a heart on the window
I am entering the security code for the building
I am fast-forwarding the tape in the stereo
I am following the line of text as I read
I am hitting escape on the keyboard
I am liking my friend's status on facebook
I am loading the weather app on my i-phone
I am navigating the touchscreen display
I am opening the DVD player on my laptop
I am opting for A5 on the photocopier
I am ordering a latte from the coffee machine
I am picking my nose in the lecture
I am pressing send on the email
I am programming the new burglar alarm
I am pushing my glasses up my nose
I am putting in my contact lens
I am requesting the elevator on the top floor
I am ringing the doorbell at my auntie's house
I am saving the taxi number on my phone
I am scratching inside my ear
I am scrolling down through the pdf
I am seeing whether the piano is tuned
I am setting the microwave to defrost
I am sounding the bell to stop the bus
I am spraying air-freshener in the kitchen
I am switching off the lamp at the plug socket
I am taking a photograph on the monument
I am turning on my laptop in the library
I am typing my pin number at the cashpoint
I am updating the central heating timer
I am writing my name in the sand

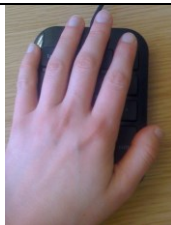
PALM sentences

I am admiring my new engagement ring
I am blaring my horn in the traffic
I am bouncing a basketball in the gym
I am checking if my housemate has a fever
I am clasping the baby tightly to my chest
I am cleaning the mirror in the hallway
I am clinging onto my hat in the blizzard
I am clutching my stomach after the meal
I am covering my eyes at the bright light
I am ensuring the wallpaper is smooth
I am feeling my pregnant sister's stomach
I am flattening creases in the tablecloth
I am giving my brother a high five
I am greeting my friend across the street
I am having my fortune lines read
I am playing the tambourine in the concert
I am pledging the oath of allegiance
I am plumping up the pillow in the hospital bed
I am protecting my face as I fall
I am reaching for the book on the shelf
I am receiving the change from the shopkeeper
I am removing the smears from the window
I am ruffling my dog's head after the walk
I am saluting the general in the parade
I am serving in the volleyball championship
I am shaking my boss's hand after the meeting
I am shielding my eyes from the sun
I am slamming shut the car door
I am slathering suntan lotion onto my neck
I am smacking my boyfriend in the face
I am smoothing my hairstyle down in the bathroom
I am squashing the hamburger before biting it
I am stifling a yawn in the seminar
I am stopping a goal in the penalty shootout
I am straightening out the bedspread
I am stroking my neighbour's cat
I am swatting a fly away from my face
I am swearing on the bible to tell the truth
I am whispering to my niece behind my hand

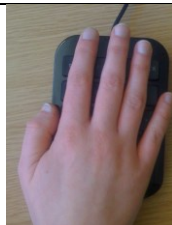
FILLER sentences

| | |
|---|---|
| <i>I am activating the alligator for my uncle</i> | <i>I am hammering insider the tennis ball</i> |
| <i>I am beating the picture frame into a mousse</i> | <i>I am hanging the lamppost outside the flat</i> |
| <i>I am beginning through the light socket</i> | <i>I am hiding the lighthouse by the sea</i> |
| <i>I am bending the shirt in the garage</i> | <i>I am ironing the boiler in the airing cupboard</i> |
| <i>I am breaking the custard into a bowl</i> | <i>I am joining the goldfish and the worktop</i> |
| <i>I am bringing the housing estate to the meeting</i> | <i>I am knitting the taxi on the road</i> |
| <i>I am burning the drinking water in the canteen</i> | <i>I am knocking the postage stamp in the tea</i> |
| <i>I am burying the paperclips in the drawer</i> | <i>I am knitting the bungalow doorstep</i> |
| <i>I am carrying the ship in a picnic basket</i> | <i>I am laying the bed in the toilet</i> |
| <i>I am catching the space shuttle in a net</i> | <i>I am leading the milk in the fridge</i> |
| <i>I am cheering the lemonade on the mat</i> | <i>I am lighting the hosepipe in the greenhouse</i> |
| <i>I am chopping a saucer for my neighbour</i> | <i>I am loading the lorry on the notebook</i> |
| <i>I am conducting a bottle of orange juice</i> | <i>I am melting the library at the meeting</i> |
| <i>I am counting the ladder on the windows</i> | <i>I am messing a carrot from the greengrocer's</i> |
| <i>I am cracking the syrup in the cupboard</i> | <i>I am mining my Walkman in the theatre</i> |
| <i>I am crashing the blossom on the trees</i> | <i>I am mowing the scientific calculator</i> |
| <i>I am crunching the mobile phone into a ball</i> | <i>I am offering the milkshake to the wall</i> |
| <i>I am cutting the cupboard with a cake fork</i> | <i>I am operating on the rusty mail</i> |
| <i>I am digging the shoe on the pavement</i> | <i>I am packing the Eiffel Tower into the train</i> |
| <i>I am dragging the puddle after the rain</i> | <i>I am parking the blanket on the motorway</i> |
| <i>I am draining the whiskey on the sunbed</i> | <i>I am peeling the kitchen with a knife</i> |
| <i>I am dressing the saucepan under the carpet</i> | <i>I am piling the pineapples in the thimble</i> |
| <i>I am drilling the paper by the motorway</i> | <i>I am planning the vacuum cleaner for dinner</i> |
| <i>I am drinking the biro for desert</i> | <i>I am pouring a wardrobe down the drain</i> |
| <i>I am dripping a ceiling fan on the floor</i> | <i>I am practicing the television remote control</i> |
| <i>I am driving the antique grandfather clock</i> | <i>I am raising the tin of beans under the oven</i> |
| <i>I am drumming the best-selling cookery book</i> | <i>I am raking the lions on the front lawn</i> |
| <i>I am dusting the chicken with a mop</i> | <i>I am reading a spoon in the nursery</i> |
| <i>I am feeding the moustache in the mirror</i> | <i>I am recording the chopping board speak</i> |
| <i>I am filing the macaroon in the documents folder</i> | <i>I am repairing the strawberry ice cream</i> |
| <i>I am filling the widescreen television</i> | <i>I am riding an apple by the canal</i> |
| <i>I am fisting the button onto the shirt</i> | <i>I am rippling a typewriter under my arm</i> |
| <i>I am firing the spaniel with pinewood</i> | <i>I am sawing the pond behind the house</i> |
| <i>I am fishing the cushion with blue cement</i> | <i>I am scraping the furniture on the pizza</i> |
| <i>I am flapping the velvet three piece suite</i> | <i>I am screwing the orange into the lounge</i> |
| <i>I am flying in the coffee table</i> | <i>I am sewing the juice into a cocktail</i> |
| <i>I am folding the fridge into three pieces</i> | <i>I am snapping the scarf into chocolate</i> |
| <i>I am freezing my colleagues who lost her wallet</i> | <i>I am threading the door handle on the dew</i> |
| <i>I am grading the police station on its essay</i> | <i>I am tightening the coin into the freezer</i> |
| <i>I am grinding the tablecloth into the coffee pot</i> | <i>I am watering my socks in the living room</i> |

ACTION Images



PALM-left



PALM-right



POKE-left



POKE-right